

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 15:02:05 ; Search time 532 Seconds
(without alignments)
4128.393 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVVEVAGSMWAQERDIIIM.....TSGPCLGHLHSSQFPFPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9267905 segs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abses/ABSSWEB_spool/US10071838/runat_05042006_141923_5354/app_query.fasta_1
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=trpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abs06h
-USER=US10071838 @CGN_1_1_541@runat_05042006_141923_5354 -NCPU=6 -ICPU=3
-NO MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGADOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq1:*
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq:*
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq1:*
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq2:*
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq3:*
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq:*
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq4:*
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	676	22.7	378	6	US-09-925-065A-506683
C 2	667	22.4	378	6	US-09-925-065A-506682

3	437	14.6	10490	14	US-11-011-332A-74	Sequence 74, Appl
4	380	12.7	1052	6	US-09-925-065A-16910	Sequence 16910, A
5	380	12.7	1052	6	US-09-925-065A-16911	Sequence 16911, A
6	380	12.7	1052	9	US-10-301-480-118147	Sequence 118147, A
7	380	12.7	1052	9	US-10-301-480-118148	Sequence 118148, A
8	380	12.7	1052	10	US-10-301-480-731556	Sequence 731556, A
9	380	12.7	1052	10	US-10-301-480-731557	Sequence 731557, A
10	367	12.3	1294	6	US-09-925-065A-55408	Sequence 55408, A
11	367	12.3	1294	6	US-09-925-065A-55409	Sequence 55409, A
12	367	12.3	1294	9	US-10-301-480-156646	Sequence 156646, A
13	367	12.3	1294	9	US-10-301-480-156647	Sequence 156647, A
14	367	12.3	1294	10	US-10-301-480-770055	Sequence 770055, A
15	367	12.3	1294	10	US-10-301-480-770056	Sequence 770056, A
16	348.5	11.7	1385	11	US-11-096-568A-23142	Sequence 23142, A
17	313.5	10.5	528	6	US-09-925-065A-534386	Sequence 534386, A
18	311	10.4	1820	11	US-11-096-568A-6617	Sequence 6617, Ap
19	297.5	10.0	804	6	US-09-925-065A-62926	Sequence 62926, A
20	297.5	10.0	804	9	US-10-301-480-164164	Sequence 164164, A
21	297.5	10.0	804	10	US-10-301-480-777573	Sequence 777573, A
22	293.5	9.8	601	9	US-10-301-480-30359	Sequence 30359, A
23	293.5	9.8	601	10	US-10-301-480-643768	Sequence 643768, A
24	292	9.8	2391	6	US-09-925-065A-69533	Sequence 69533, A
25	292	9.8	2391	6	US-09-925-065A-69535	Sequence 69535, A
26	292	9.8	2391	9	US-10-301-480-170772	Sequence 170772, A
27	292	9.8	2391	9	US-10-301-480-170774	Sequence 170774, A
28	292	9.8	2391	10	US-10-301-480-784181	Sequence 784181, A
29	292	9.8	2391	10	US-10-301-480-784183	Sequence 784183, A
30	290.5	9.7	2391	6	US-09-925-065A-69534	Sequence 69534, A
31	290.5	9.7	2391	9	US-10-301-480-170773	Sequence 170773, A
32	290.5	9.7	2391	10	US-10-301-480-784182	Sequence 784182, A
33	268	9.0	1870	14	US-11-124-368A-38	Sequence 38, Appl
34	256.5	8.6	578	14	US-11-128-061-2085	Sequence 2085, Ap
35	256.5	8.6	578	14	US-11-128-061-5727	Sequence 5727, Ap
36	256.5	8.6	578	14	US-11-128-049-2085	Sequence 2085, Ap
37	256.5	8.6	578	14	US-11-128-049-5727	Sequence 5727, Ap
38	253	8.5	2834	14	US-11-000-688-949	Sequence 949, Appl
39	248	8.3	3431	14	US-11-169-041-16	Sequence 16, Appl
40	241	8.1	10490	14	US-11-011-332A-86	Sequence 86, Appl
41	236.5	7.9	1902	9	US-10-932-182A-78197	Sequence 78197, A
42	236.5	7.9	1902	9	US-10-932-182A-78197	Sequence 78197, A
43	236	7.9	879	6	US-09-925-065A-9583	Sequence 9583, Ap
44	236	7.9	879	9	US-10-301-480-110820	Sequence 110820, A
45	236	7.9	879	10	US-10-301-480-724229	Sequence 724229, A

ALIGNMENTS

RESULT 1
US-09-925-065A-506683/c
; Sequence 506683, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 506683
; LENGTH: 378
; TYPE: DNA

ORGANISM: Homo sapiens
US-09-925-065A-506683

Alignment Scores:
Pred. No.: 6,52e-33 Length: 378
Score: 676.00 Matches: 119
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 22.7% Indels: 0
DB: Gaps: 0

US-10-071-838-2 (1-549) x US-09-925-065A-506683 (1-378)

QY 361 AAlaYSPROGLUGInGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 357 GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCRCCTTCACGTCGGCGGAAG 298
QY 381 ThrLeuCySlySGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 297 ACCCTCTGCAGAGGGGACAGGACGGCCCTCCAGGCCACAGCCGGTTCCCGCGGCC 238
QY 401 ILeTSPeAlaSerProProArgAlaProArgSerSerThrProCySPROGLyGlyAla 420
Db 237 ATTTGGTCAGCTTCCCGCCACAGGCGACCTGCTTCCACACACCTGCTGCTGGGCT 178
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 177 GTCCGGAGACACCTACCTCTGTGGCACTCAGGCTGTGCCAGCCCGCTGCTCAG 118
QY 441 GlyGlyProGlnGlySerTyrParGpHelaGlnTyrPasnSerMetProArgLeuProThr 460
Db 117 GGAGGACCTCAGGGTCTCTGGAGATTCTCGAGTGAACCTCATGCCCGCTCCCAACG 58
QY 461 AspLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyr 479
Db 57 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGCTGG 1

RESULT 2
US-09-925-065A-506682/c
Sequence 506682, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 506682
LENGTH: 378
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-506682

Alignment Scores:
Pred. No.: 2,32e-32 Length: 378
Score: 667.00 Matches: 118
Percent Similarity: 99.2% Conservative: 0
Best Local Similarity: 99.2% Mismatches: 1
Query Match: 22.4% Indels: 0
DB: Gaps: 0

US-10-071-838-2 (1-549) x US-09-925-065A-506682 (1-378)

QY 361 AAlaYSPROGLUGInGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 357 GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCRCCTTCACGTCGGCGGAAG 298
QY 381 ThrLeuCySlySGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 297 ACCCTCTGCAGAGGGGACAGGACGGCCCTCCAGGCCACAGCCGGTTCCCGCGGCC 238
QY 401 ILeTSPeAlaSerProProArgAlaProArgSerSerThrProCySPROGLyGlyAla 420
Db 237 ATTTGGTCAGCTTCCCGCCACAGGCGACCTGCTTCCACACACCTGCTGCTGGGCT 178
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 177 GTCCGGAGACACCTACCTCTGTGGCACTCAGGCTGTGCCAGCCCGCTGCTCAG 118
QY 441 GlyGlyProGlnGlySerTyrParGpHelaGlnTyrPasnSerMetProArgLeuProThr 460
Db 117 GGAGGACCTCAGGGTCTCTGGAGATTCTCGAGTGAACCTCATGCCCGCTCCCAACG 58
QY 461 AspLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyr 479
Db 57 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGCTGG 1

RESULT 3
US-11-011-332A-74
Sequence 74, Application US/11011332A
Publication No. US20060024684A1
GENERAL INFORMATION:
APPLICANT: Foekens, John
APPLICANT: Harbeck, Nadia
APPLICANT: Koenig, Thomas
APPLICANT: Maier, Sabine
APPLICANT: Martens, John
APPLICANT: Model, Fabian
APPLICANT: Nimmerich, Inko
APPLICANT: Rujan, Tamas
APPLICANT: Schmitz, Manfred
APPLICANT: Lesche, Ralf
APPLICANT: Dietrich, Dimo
APPLICANT: Mueller, Volkmart
APPLICANT: Kluth, Antje
APPLICANT: Schwoppe, Ina
APPLICANT: Hartmann, Oliver
APPLICANT: Adorjan, Peter
TITLE OF INVENTION: PROGNOSTIC MARKERS FOR PREDICTION OF TREATMENT RESPONSE AND/OR S
TITLE OF INVENTION: BREAST CELL PROLIFERATIVE DISORDER PATIENTS
FILE REFERENCE: 47675-99
CURRENT APPLICATION NUMBER: US/11/011, 332A
CURRENT FILING DATE: 2004-12-13
PRIOR APPLICATION NUMBER: US 10/517, 741
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: PCT/EP2003/010881
PRIOR FILING DATE: 2003-10-01
PRIOR APPLICATION NUMBER: DE 10245779.4
PRIOR FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: DE 10300096.8
PRIOR FILING DATE: 2003-01-07
PRIOR APPLICATION NUMBER: DE 10317955.0
PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: PCT/EP2004/014170
PRIOR FILING DATE: 2004-12-13
PRIOR APPLICATION NUMBER: EP 03090432.0
PRIOR FILING DATE: 2003-12-11
PRIOR APPLICATION NUMBER: EP 04090041.7
PRIOR FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: EP 04090380.9
PRIOR FILING DATE: 2004-09-30
PRIOR APPLICATION NUMBER: EP 04090127.4
PRIOR FILING DATE: 2004-04-01

; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 158
 ; SEQ ID NO 74
 ; LENGTH: 10490
 ; TYPE: DNA
 ; ORGANISM: Homo Sapiens
 ; US-11-011-332A-74

Alignment Scores:

Pred. No.:	4,71e-17	Length:	10490
Score:	437.00	Matches:	254
Percent Similarity:	15.4%	Conservative:	4
Best Local Similarity:	15.2%	Mismatches:	20
Query Match:	14.6%	Indels:	1401
DB:	14	Gaps:	9

US-10-071-838-2 (1-549) x US-11-011-332A-74 (1-10490)

QY	24	lysGlyHisArgAlaGlyLeuProGluAspLysGly-----ProLysProPheArgSer	41
DB	5332	CAGGGCTCCAGAGCCCAAGACCAGCATCCGGCGGCTCTGGGAAGCCTGGCAGCTCC	5391
QY	42	TyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeuThr	61
DB	5392	GCTAACTCCAACATGCTCATTTG-----	5415
QY	62	AlaArgGluAlaLys-GlnIleArgArgGluIleSerArgLysSerLysTrpValAspMe	81
DB	5416	-----ACAGCAAAATTCGGCGGAGATCAGCCGAAGAGCAAGTGGTGATAT	5463
QY	81	tLeuGlyAspTrpGlyLysTrpLysSerSerArgLysLeu-----	94
DB	5464	GCTGGAGACTGGGAATAATACAAAGACAGAGAAAGT-AACTGTGAGAGGAGAAAGC	5522
QY	94	-----	94
DB	5523	ACTCTGTGAGAGACAGGGACAGGACCCATGCTGTGGCTTGCAACCATCAGCTCTC	5582
QY	94	-----	94
DB	5583	AGAGGGTGGCGGCACACTGTCTCCGCCAGAGACTGCAGGCCCTGTCGCGAGATTTC	5642
QY	94	-----	94
DB	5643	TGCCTATTGTGCAAGCGTCACTTGCAAGGAGGAATCTGAATCTAGGGCTGGACTAC	5702
QY	94	-----	94
DB	5703	CCGGAGCTCAAGGCTAGGATGCCCTGTGTGACTGAAGGAAGAAAGTTCAAGTCAGA	5762
QY	94	-----	94
DB	5763	GTTTCGACTCTGAGTGTCCATCCACTTTTCAGTCCGGGAAGGAGACCTGTCCACG	5822
QY	94	-----	94
DB	5823	TTGATCTCACTCTACTGAGGAATCATGGGGCCAAACCACAATTTCAGAAATCCCGG	5882
QY	94	-----	94
DB	5883	GCTTGTCTCTCACTGGGGTACCCCCGTGGCTGTGACACAGATCGTTTCTGCCACA	5942
QY	95	-----IleAspArgAlaTyrIleGlyMetProMetAsnIleArgGlyProMetTrpSerVa	113
DB	5943	GCTCATGATCGAGCGTACAGGGAATGCCCATGAACATCCGGGCGCGATGTGTCAGT	6002
QY	113	IleuLeuAsnIleGlyMetLysLeuLysAsnProGlyArgTyrGlnIle-----	130
DB	6003	CCTCTGAACATTGAGGAATGAAGTGAAGAAACCCCGGAAGATACCAAGT-ACGCTCAG	6061
QY	130	-----	130
DB	6062	CCAGAGCAACAACAAGACAGAGCCGTGTGGGGCCAGGTCTCCAGCTGAGGGAAGC	6121

QY	130	-----	130
DB	6122	TCAAGACCACTGGGAGCTGGGGGTGAAGTCAAGATGAACCCCTGGCACAGATGCT	6181
QY	130	-----	130
DB	6182	GACACAGTCAACACAGCAAACTGAGCTGTGTGACCCCTCGCTTCACTAACAAGCC	6241
QY	130	-----	130
DB	6242	AAATGCAAGCTTTCTGCAAGAAAGAACTTCTCTGTCTCTCTCTCCGAAAGTGTGAC	6301
QY	130	-----	130
DB	6302	TGTGGCTGACTGCCACTGGGGGAGAGAGTCTTCCATCTGTCTGAGACTGCTTCTCC	6361
QY	131	-----MetLysGlyLysGlyLysArgSerSerGlyHisIleG1	143
DB	6362	TCTTGCCCTGCTTACAGATCATGAAGAGAGGCAAGGTCATCTGAGCACATCCA	6421
QY	143	naArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTy	163
DB	6422	GCGCATGCAACGGGACGTAAGCGGCATTAAGAAAGCATATATCTTCAAGGATCGATA	6481
QY	163	rglyThrLys-----	166
DB	6482	CGGAACCAAGTAAGCTTACGGGAGCCACAGGTTCCACAGAGATGGGGTGAATGAGAG	6541
QY	166	-----	166
DB	6542	GATGGGGCTTCCCCGAGAGAAAGCCAGGTCACCCAGAGGAGTACACAGCTGCCAA	6601
QY	166	-----	166
DB	6602	GAGCTTCCGGCCAGGAGAGAGCCGCCACCATGAACCGAGCACTCCCTGTTCGAAG	6661
QY	166	-----	166
DB	6662	CCCTGGCCAGACTGMAACATGTGGGCCAGAAACCAAGAGATCTTGAGAGATGGAAG	6721
QY	166	-----	166
DB	6722	GCAGCAACAATAATCATGACATGTGTAAGGCTGTCTCCCTGACCCATGGGGAACCAT	6781
QY	166	-----	166
DB	6782	GGTAGACCCACGGGAGGTGGCAGATAGAGGGCCCATGAGCCCCCCAGGCAACAGT	6841
QY	166	-----	166
DB	6842	GACAGACCAATGCTGGAGAAATTAGGGTCTTGAAACTTCATCCAGTCCGCTGGG	6901
QY	166	-----	166
DB	6902	AACATGACATGGCACAGCCAGCTTGGCAGCCCGTTGGCAGTGGCTCACAAAGCTCGATG	6961
QY	166	-----	166
DB	6962	GACTGAACACACATCCCAAGTGTACAGATATTGAACCACTGATTTGCAAACTGA	7021
QY	166	-----	166
DB	7022	CATCCATGAACAACAGATGCCAGTTCACTGCTTGACTCTCTGTCACTCACACGGA	7081
QY	166	-----	166
DB	7082	GCTTGGGGGACGGCTTCAACACGGGATGGGAGAGCAAGGCTGTCTCTCTTCAAA	7141
QY	166	-----	166
DB	7142	CGAAGACCCAGTGAGAAAGGGAACGAGCCGGTGTATGCCGACGAACGTGGGTGATC	7201

QY	166	-----	166
Db	7202	CTAGATGATTTTGCTGAGGACAGAAAGCCAGACCAATTAAGCTACCAAGTAGATTCC	7261
QY	166	-----	166
Db	7262	CATTCCTAGGCCATTCTGAAAAAGCCAAACCAAGGACTGAGAAGCAGTCTGGGTGC	7321
QY	166	-----	166
Db	7322	CAGGGCTGACCGATCGGGGAGAGGCTGGGTGATAGGGCCACCTGGAGACTTGAGG	7381
QY	166	-----	166
Db	7382	ATGAAGAGTGCCTCCAGAGAGGGGCTGAGCGGTGCCGGGAGACTCTGCACATTGGTTT	7441
QY	166	-----	166
Db	7442	GGAACCGTGAGGAATGTACACCCACAGACTGAACCTGGCGTGTGCAAACTGAAAAA	7501
QY	166	-----	166
Db	7502	AAAAAAAAAATCATTCAGAGTGAAGAATCAGGCAAGTCACTGTACAACTGGGCTATT	7561
QY	166	-----	166
Db	7562	TGCATGTCAAGATGTGGATTTTACTGAACAATTCTTCAAGATCTCAGGCCCTGAGA	7621
QY	166	-----	166
Db	7622	GCTCACTGCTTATCTGTGTGAACAATCTGAACCTGAATGGGATTGCTTAGGCTTTGT	7681
QY	166	-----	166
Db	7682	AGACAAAGTGAATTAACAACATCTGCACAAAACAAAACCAAGCCCCCTTCTCTGTTTC	7741
QY	167	-----GlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluTyrAsnPro-----	182
Db	7742	CTAGGCAAGCGGAATCTCCACATCTCTCGGCATATGAGGAGTCAACCCGGTGAAGTA	7801
QY	182	-----	182
Db	7802	TTCCCCGCAGTGAGTTCCCCGGCCATATTTCCATATTGACAGAGTGGGTCTGTGG	7861
QY	182	-----	182
Db	7862	GAGTGTGCTTCTTTTAAAGTTAGTATTGTGACCCACCAAGATATAGAGGTAGGA	7921
QY	182	-----	182
Db	7922	TGTCAAGTCAACCGCTGGCATAACTCCAAAGAGGGGTGTTCTCAAGGGGTCAAGCTG	7981
QY	182	-----	182
Db	7982	AGACACAGAGAGTCAAGGCTTGACTCTGTGTCTCACTGGGCTGACCAACCACTTCTC	8041
QY	182	-----	182
Db	8042	AGACAAGAATAAGACGCTCTCTCGGGCTGCCCAAAAGCCACAGAGCTTGGCAGCAT	8101
QY	182	-----	182
Db	8102	CGCACACAGAGTGTCTATCAGACAGACATTTTGACAAGGTGCTGAAGTGCCTGATGA	8161
QY	182	-----	182
Db	8162	CTTGCTCTTGTATGAATAATGATGTGATCTGAGGAAGCTCTTTTTCAGAGGAAGCC	8221
QY	182	-----	182
Db	8222	TCTCCTTCAGAGGAAGCTCTCCAGTCACTCTGCCCTCTCCAATGACATGATCCTCCC	8281
QY	183	-----GluV	184

Db	8282	AGTGACCTCAGCCCTCCCAAGTGAATGCTCTTCCATGTGACTCTGCTCTTGACAGAGG	8341
QY	184	AlGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyrLeuProGluG	204
Db	8342	TGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTATCTCTGAGG	8401
QY	204	IuAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGln----	222
Db	8402	AGGATGATTTCTGGCACTGTGTGACGTCTGTGCCAGTGAAGGACACTCCCTGCA-GGGT	8460
QY	222	-----	222
Db	8461	AAGTGAAGCTGCCCCGGGACCTCTGCAGCCAGACTGGGATGGCCACCTGGCCA	8520
QY	222	-----	222
Db	8521	GGTGATCAGCTTTTCAAGCCAAAGCACTCTGTGTGTGCCAGCTTGTGGAGACTTT	8580
QY	222	-----	222
Db	8581	AGGATGTCTGTGAGGTTCCCAAGAGTCCACGGGTGACCCCAAAAGCCAAATCAG	8640
QY	222	-----	222
Db	8641	ACGCTTCTATCCCATCAGCAGAGGGCATCTATCTCCCCGTGGCACTCTGTGTC	8700
QY	222	-----	222
Db	8701	CTGAGCCAGCCCTCCGGCTGTGATTCTGTGACGTGACTCTCCCTCTGAGAGTCC	8760
QY	222	-----	222
Db	8761	TCCTGCCCTCAGCTGCCCGGCTCTGTCTGTCCATCGGTGCCACGAATGGCGGACCA	8820
QY	222	-----	222
Db	8821	GCCCAAGTGGCAGCATCTCCCAATCCCTGTTCCTGTGGCCGACCCCACTACAGAGAT	8880
QY	222	-----	222
Db	8881	GACCGGAAGCCCAAGCCCAACCAATTCGGGCCACCTGTGTGGCTGAAGTCAAGC	8940
QY	222	-----	222
Db	8941	TTGCCCTTTTTCACCTGTGGCCAGAGGCTCCAGGGGAACCTCAGCCAGGCTCCAGG	9000
QY	222	-----	222
Db	9001	GAATGTCCGCCCCCACTCCCAAGGTAAAGCCGATGTTGGGTCAACAGATGGAG	9060
QY	222	-----	222
Db	9061	GGTGAGGCTTGGGTTTGGGGCTCTCCAGCTGCCAGCTCTTGACGTGATGGCT	9120
QY	223	-----GlyPheHis	226
Db	9121	CCACATCTTGGGGGAAGCTCTGATTTCATGATGGCTGGGGCTTCTCAGGATTTCA	9180
QY	226	erProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnIleValIleThrS	246
Db	9181	GCCCAATGTGGGGACCGTCCAGGGGCTCCAAGACCAAGAGCATGTGTAGCCACGT	9240
QY	246	erGlnProIysThrMetGly-----	252
Db	9241	CACAAACCAAGACATGGGCATCAGGTGAGTTATGTCCCTCAGCTCTTCCAGAGG	9300
QY	252	-----	252
Db	9301	CCCTGCTCCCGTGGGCTGTAGAGCAGGGGGCTGAGCCCTCGTGGGCTGTGAC	9360
QY	252	-----	252

Db 9361 TGGCTAGTCCCAGCCAGGGCCTGACCTGGGACGTGGGTTCTTCATGGGCTGGAGTTG 9420
QY 252 ----- 252
Db 9421 GTTTCCTTCTGCTGAGAGAGACAGAGGACAGGATGGGGCCAGCTCCCGACA 9480
QY 252 ----- 252
Db 9481 GCAGGGCAAGGCACTGTCTCCACCGGAGTGTGGGAAGTGACAGTGTGTGGGAGC 9540
QY 252 ----- 252
Db 9541 TCTGACACCCCGCAGTGTCTGCACTAGGGGAGGGTCTTACAGGGCCCTGGAAGAGG 9600
QY 253 -----HisGlnAspLysLysA 258
Db 9601 AGCTTTTAGGGCAGCCCACTGCTGAGCACTCTGTTGCTTCCATCAGACAGAAAG 9660
QY 258 sPLeuCy sGlyInCy sSerProLeuGlyCy sLeuIleArgIleLeuIleAsp-Gly--- 276
Db 9661 ATCTATGTGGGCACTGTTCCTGAGGCTGCTCATCCGATATTGATTGACGGGTA 9720
QY 276 ----- 276
Db 9721 GGAGCATAGGAGACCCCTGGCTCAGGGACTTCTTGCCCTGCAGTGCCTGCTTCCCC 9780
QY 276 ----- 276
Db 9781 AGCCCGGGGTCTGCTCACTCCACGCCACAGAGGCTCAGGGGGTCCCAAGAGACA 9840
QY 276 ----- 276
Db 9841 CACAAGCAAAACCTCTGCCCCAAGGGGTCTATCCAGGGCCATGCTGGGCTCAGGCC 9900
QY 276 ----- 276
Db 9901 CAGCCTCATGGGAGACTGGGCCAGACCAGCTTGAGAGGGCTCAGGGAAGCCTCAAGC 9960
QY 276 ----- 276
Db 9961 CCTGGCAAGCCCTCTCTCCAGAGCCATCCCACTCAATGAGTCCCCCATGAG 10020
QY 276 ----- 276
Db 10021 GAGCTTCAAGACCTTGTCTGACCCAGCGTCTGAGGGCTCAGGCGAACCTCATGGGAA 10080
QY 276 ----- 276
Db 10081 GGTCACTGACTTGGAGACTGAAGCCCCAGTGTGCGAGCTGAGCCACCGCCCAAGC 10140
QY 277 -----IleSerLeuGlyLeuThrLeu 283
Db 10141 TGAAGAGACCAAGTCTTTCAACACTGCTGTCCCAACAGATCTCTCGGGCTCACCTG 10200
QY 284 ArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIleThrArgIle 303
Db 10201 CGCCTGTGGACGTGTATCTGCTAGAGCGCAAGCGCTGATGCCGATAACAAGATC 10260
QY 304 AlaPheLysValGlnGlnLysArgLeuThrLysThrSerArg 317
Db 10261 GCCTTTAAGGTTACAGCAGAGTAAGTTACGTGTGCCAGCGG 10302
RESULT 4
US-09-925-065A-16910
; Sequence 16910, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16910
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-16910
Alignment Scores:
Pred. No.: 2.05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 6 Gaps: 13
US-10-071-838-2 (1-549) x US-09-925-065A-16910 (1-1052)
QY 277 IleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAla 296
Db 7 ATCTCTCGGGCTCACCCCTGCGCTGTGGAGCGTGTATCTGTGTAAGGCGCAAGCGG 66
QY 297 LeuMetProIleThrArgIleAlaPheLysValGlnGln 309
Db 67 TTGATGCCGATAACAAGATCGCTTAAAGTTCAAGAGTAAGTCTACGTGTGCCAG 126
QY 309 ----- 309
Db 127 CGGGCCTGGGAGCCCTGGGGTCAAGCCCGACTGGCCCGAGGCAAGCTTCTCACT 186
QY 309 ----- 309
Db 187 GTCTCATGATCCKCTGTTCTGGCCAGAGGAGTCTGGCCAGGTGGGCTGGCAGGAC 246
QY 309 ----- 309
Db 247 ACTGTACACCGAGCCCATCCCAATGACCCAGATGAAGTCAGAGTGTGTAGACA 306
QY 309 ----- 309
Db 307 CTTCCTGTCCGATCGCCCCCAGCCACAGTCTCCTGTGTATATCTGACACCTGGGCT 366
QY 309 ----- 309
Db 367 GGCCCAAAAGATCCGGCACCGGCCAGTAGAGACTGAAGTGGCCACGGGATATGACT 426
QY 310 -----Lys-ArgLeuTh 313
Db 427 GTGACCATTTCCAGGTAAGTACTCCCTGGCTGATATCCACCCTGTCTCTAGAGCGCTCAC 486
QY 313 rLysThrSerArgCy sGlyProTrpAlaArgPheCy sAsnArgPheValAspThrTrpAl 333
Db 487 GAAGACGTCCAGGTGTGGCCCGTGGCCAGCTTTTGGCAACCGTTCGTTGATACCTGGGC 546
QY 333 aArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLy 353
Db 547 CAGGATGAGACACTGTGTCTCAACATCTTAGGGCTCTATGAAGAACTAACAGAAA 606
QY 353 sGlnGlyAspLeuProProPro-AlaLysProGluGlnGlySerSerAlaSerArgProv 373
Db 607 GCAGGGGACCTGCGCACCCCGAGGTG-----GGCTCCAGTGCATGTCCCT- 653
QY 373 alProAlaSerArgGlyGlyLysThrLeuCy sLysGlyAspArgGlnAlaPro-ProGly 392

```

Db      654  --CCCATGTACCCCTCTGGGGTAGTCAGTACTAGGGGA-----GTGCCCGGGA 699
QY      393  PROProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaProArgSer 412
Db      700  CCCGCAACCCCTACTACTGGGCT-----TCCTCTTACCTTTTCTTCT----- 744
QY      413  SerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnGly 432
Db      745  -----CCTCTTCTCTCTGACT-CTAAGAAAGTACAGAGGCCCAACCGCTCTCAGGGC 797
QY      433  ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrp 452
Db      798  AGCGCTCAGTGGCT-----GTATACTGG 821
QY      453  AsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheArgHisTyr 472
Db      822  ACATGCTGTGCACGC----- 836
QY      473  AspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db      837  -----AGAGGGGGATGTGGCAAGA----- 857
QY      493  TrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db      858  -----CCCTCCACAAGCCCCCTCCCACTTTCACGGTGTCTCCCTCTCC 902
QY      513  SerAspGlnGly-----ThrProPheArgAlaArgAspGlu 524
Db      903  CCCTCGAGGGCCCTCCAAGTTACTAGACGAGCCAGACCCCATTTGTGGAGAGACCCCGCC 962
QY      525  GlnProCys--AlaProThrSer-----GlyProCysLeuCysGlyLeuH 539
Db      963  CCTCCCTGCAGACACCAAGCCTCAGAGAGCAGAGAGGCCCTCACTC-----CTGC 1016
QY      539  IseLeuGluSerSerGlnPheProPro 547
Db      1017  ACGCT-----CCTCCA 1027
```

```

RESULT 5
US-09-925-065A-16911
; Sequence 16911, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16911
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-16911
```

```

Alignment Scores:
Pred. NO.:      2.05e-14      Length:      1052
Score:          380.00      Matches:      133
Percent Similarity: 34.7%      Conservative: 10
Best Local Similarity: 32.3%      Mismatches: 60
```

```

Query Match:      12.7%      Indels:      210
DB:                6          Gaps:      13
US-10-071-838-2 (1-549) x US-09-925-065A-16911 (1-1052)
QY      277  ILeSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAla 296
Db      7    ATCTCTCTGGGGCTCACCCCTGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAAGGGC 66
QY      297  LeuMetProIleThrArgIleAlaPheLysValGlnGln----- 309
Db      67  TTGATGCCGATTAACAAGAAATCGCTTTAAGGTTCAAGCAGAGTAAGTCTACGTGTGCCAG 126
QY      309  ----- 309
Db      127  CGGGCCTGGGGAGCCCTGGGGTCAAGCCGACTGGCCGAGGGCAGCTTCTCAGACT 186
QY      309  ----- 309
Db      187  GTCCTCATGATCCTCTGTTCTGGCCAGAGGAGTCTGCGCCAGGTGGCTGGGCAAGAC 246
QY      309  ----- 309
Db      247  ACTGTGACACCGAGCCCATCCCCACATGACCCAGATGAAGTCGAGAGTGTGTGACGA 306
QY      309  ----- 309
Db      307  CTTCCCTGTCCGGATCGCCCCCAGCCACAGTCTCTGTGTATATCTGGACACCTGGGGT 366
QY      309  ----- 309
Db      367  GGCACAAAGAGATCCGGACCGCCAGTAGAGACTGAAGTGCCACGGATATGAGCT 426
QY      310  ----- 486
Db      427  GTGACCATTCAGGTAACCTCCCTGTGCTGATATCCACCCTGTCCCTAGAGCGCTCAC 486
QY      313  rLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAl 333
Db      487  GAAGACGTCCAGGTGTGGCCCGTGGGCAAGTCTTTTGCAACCGGTTCGTTGATACCTGGGC 546
QY      333  aArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLy 353
Db      547  CAGGATGAGGACACTGTGCTCAAGCATCTTAGGGCCTTATGAAGAACTAACAGAAA 606
QY      353  sGlnGlyAspLeuProProPro--AlaLysProGluGlnGlySerSerAlaSerArgProV 373
Db      607  GCAGGGGAGCTGCACCCCCAGGTG-----GGCTCCAGTGCCATGTCCCT- 653
QY      373  aProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPro--ProGly 392
Db      654  --CCCATGTACCCCTCTGGGTAGTCAGTAGAGGA-----GTGCCGGGA 699
QY      393  PROProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaProArgSer 412
Db      700  CCCGCAACCCCTACTACTGGGCT-----TCCTCTTCACTTTTCTTCT----- 744
QY      413  SerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnGly 432
Db      745  -----CCTCTTCTCTCTGACT-CTAAGAAAGTACAGAGGCCCAACCGCTCTCAGGGC 797
QY      433  ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrp 452
Db      798  AGCGCTCAGTGGCT-----GTATACTGG 821
QY      453  AsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheArgHisTyr 472
Db      822  ACATGCTGTGCACGC----- 836
QY      473  AspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db      837  -----AGAGGGGGATGTGGCAAGA----- 857
```

QY 493 TrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db 858 -----CCCTCCAAACAAGCCCCCTCCACTTTCACGGGTGTCTCCCTCTCC 902
QY 513 SerAspGlnGly-----ThrProPheArgAlaArgAspGlu 524
Db 903 CCCTCGCAGAGGCCCTCCAAAGTACTAGACGAGCCCAAGACCATTTGTGGAGAGACCCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCysLeuCyseGlyLeuH 539
Db 963 CCTCCCTGCAGACACCCACAGCCTCAGAGACGACGAGGCCCCCTCACTC-----CTGC 1016
QY 539 IsLeuGluSerSerGlnPheProPro 547
Db 1017 ACGCT-----CCTCCA 1027

RESULT 6
US-10-301-480-118147
; Sequence 118147, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 118147
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-118147

Alignment Scores:
Pred. No.: 2.05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 9 Gaps: 13

US-10-071-838-2 (1-549) x US-10-301-480-118147 (1-1052)

QY 277 ILeSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAla 296
Db 7 ATCTCTCTCGGGCTCACCCCTCGCTGTGGACGTGTATCTGTAGAAGCGGAACAGCGG 66
QY 297 LeuMetProIleThrArgIleAlaPheLysValGlnGln----- 309
Db 67 TTGATGCCGATTAACAAGATCGCTTTTAAGTTCAAGCAAGTAAGTCTACGTGCGCCAG 126
QY 309 ----- 309
Db 127 CGGGGCTGGGAGACCTGTGGGTGAGACCCGACTGGCCGAGGGCAGCTTCCTCACT 186
QY 309 ----- 309
Db 187 GTCTCATGATCCKTGTCTTGCGCCAGAGGAGTGTGGCCAAGTGGGCTGGGAGAGAC 246
QY 309 ----- 309
Db 247 ACTGTGACACGAGCCCATCCCCCAATGACCCAGATGAAGTCAAGAGTGTGTGAGCA 306
QY 309 ----- 309
Db 307 CTTCCCTGTCCGGATGCGCCCCCAAGCCACAGTCTCTGTGTATATCTGACACCTGGGGT 366

QY 309 ----- 309
Db 367 GGCCACAAGAAAGATCCGGCACCCGCCAGTAGAGACTGAAGTGGCCACGGATATGAGCT 426
QY 310 -----Lys-ArgLeuTh 313
Db 427 GTGACCATTTCCAGGTAAGTACTCCCTGGCCTGATATCCACCCTGTCCCTTAGAGCGCCTCAC 486
QY 313 rLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAl 333
Db 487 GAAGACGTCCAGGTGTGGCCCGTGGGACGTTTGTGCAACCGGTTGTTGATACCTGGGC 546
QY 333 aArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLy 353
Db 547 CAGGATGAGGACACTGTGCTCAAGCATCTTAGGGCCTCTATGAAGAACTAAACAAGAA 606
QY 353 sGlnGlyAspLeuProProPro-AlaLysProGluGlnGlySerSerAlaSerArgProV 373
Db 607 GCAGGGGAGCTGCCACCCCAAGGTG-----GGCTCCAGTGCCATGTCCCT- 653
QY 373 alProAlaSerArgGlyGlyLysThrLeuCyseLysGlyAspArgGlnAlaPro-ProGly 392
Db 654 --CCCATGTCAACCTCTGTGGGTAGTCAAGTAGAGGGA-----GTGCCCGGGA 699
QY 393 ProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaProArgSer 412
Db 700 CCGGCAACCTACTACTCGGGCT-----TCCTCTTCACTTTTCTTCT- 744
QY 413 SerThrProCysPProGlyGlyAlaValArgGluAspThrTyPProValGlyThrGlnGly 432
Db 745 -----CCTCTTCTCTCTGACT-CTAAGAAAGTACAGAGGCCACCGGTCTCAGGCG 797
QY 433 ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrp 452
Db 798 AGCGCTCAGTGCCT-----GTACTACTGG 821
QY 453 AsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheArgHisTy 472
Db 822 ACATGCTGTGACGC----- 836
QY 473 AspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db 837 -----AGGAGGGGATGTGGGCAAGA----- 857
QY 493 TrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db 858 -----CCCTCCAAACAAGCCCCCTCCACTTTCACGGGTGTCTCTCTCTCC 902
QY 513 SerAspGlnGly-----ThrProPheArgAlaArgAspGlu 524
Db 903 CCCTCGCAGAGGCCCTCCAAAGTACTAGACGAGCCCAAGACCATTTGTGGAGACCCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCysLeuCyseGlyLeuH 539
Db 963 CCTCCCTGCAGACACCCACAGCCTCAGAGACGACGAGGCCCCCTCACTC-----CTGC 1016
QY 539 IsLeuGluSerSerGlnPheProPro 547
Db 1017 ACGCT-----CCTCCA 1027

RESULT 7
US-10-301-480-118148
; Sequence 118148, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118148
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-118148

Alignment Scores:
Pred. No.: 2.05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 9 Gaps: 13

US-10-071-838-2 (1-549) x US-10-301-480-118148 (1-1052)

QY 277 lIeSerLeuGlYleuThrLeuArgLeuTrpAspValTyrLeuValGluglYgluGlnAla 296
Db 7 ATCTCTCTCGGGCTCACCCCTGCGCTGTGGACGTGTATCTGTGTAGAAAGCGAACAGCGC 66
QY 297 LeuMetProIleThrArgIleAlaPheIysValGlnGln----- 309
Db 67 TTGATGCCGATACAGAATCGCCTTTAAGTTCAAGACAGTAAGTCTACGTGTGCCAG 126
QY 309 ----- 309
Db 127 CCGGGCCTGGGAGCCCTGGGGTCAAGACCCGACTGGCCCGAGGGCAGCTTCTCACACT 186
QY 309 ----- 309
Db 187 GTCTCTCATGATCTCTGTCTGTCTGGCCACAGAGGAGTCTGCGCAGGTGGGCTGGCAGAGC 246
QY 309 ----- 309
Db 247 ACTGTGACACCGAGCCCATCCCCACATGACCAGATGAAAGTCGAGAGTGTGTAGACA 306
QY 309 ----- 309
Db 307 CTTCCTGTCCGATCGCCCCCAGCCACAGTCTCTGTATATCTGGACACCTGGGGT 366
QY 309 ----- 309
Db 367 GGGCCAAAAAGGATCCGGGCAACGGCCAGTAGAGACTGAAGTGGCCACGGGATATGAGCT 426
QY 310 -----Lys-ArgLeuTh 313
Db 427 GTGACCATTCACAGTAACCTCCCTGCGCTGATATCCACCCTGCTCCCTAGAGCGCTCAC 486
QY 313 rLysThrSerArgCysGlyProTrrPAlaArgPheCysAsnArgPheValAspThrTrpAl 333
Db 487 GAAAGCGTCCAGGTGTGGCCCGTGGGCACGTTTTCACACCGGTTCGTGATACCTGGGC 546
QY 333 aArgAspGluAspThrValIleuLysHisLeuArgAlaSerMetLysLysLeuThrArgLy 353
Db 547 CAGGGATGAGACACTGTGCTCAAGCATCTTAGGGCTCTATGAAGAACTAACAAGAAA 606
QY 353 sGlnGlyAspLeuProProPro-AlaLysProGlnGlnGlySerSerAlaSerArgProV 373
Db 607 GCAGGGGGACCTGCCACCCCCCAGGTG-----GGCTCCAGTGCATGTCCCT- 653
QY 373 alProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPro-ProGly 392
Db 654 --CCCATGTCAACCTCTGGGGTAGTCAGTAGTAGGGGA-----GTGCCCGGGA 699
QY 393 ProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaProArgSer 412
Db 700 CCGCGCAACCTACTACTGGGGCT-----TCCTCTTCACCTTTTCTTCT- 744
QY 413 SerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnGly 432

Db 745 -----CCTCTTCCTCCTGACT-CTAAGAAAGTACAGAGGCCACCGGTCTCAGGGC 797
QY 433 ValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrp 452
Db 798 AGCGGCTCAGTGGCT-----GTAATACTGG 821
QY 453 AsnSerMetProArgLeuProThrAspLeuAspValGluglYProTrpPheArgHisTyr 472
Db 822 ACATGCTGTGCACGC----- 836
QY 473 AspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db 837 -----AGGAGGGGAGTGTGGCAAGA----- 857
QY 493 TrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db 858 -----CCCTCCAAACAAGCCCCCTCCCACTTCCACGGTGTCTCCCTCTCC 902
QY 513 SerAspGlnGly-----ThrProPheArgAlaArgAspGlu 524
Db 903 CCCTCGCAGGGCCCTCCAGTACTAGACGAGCCACAGCCCATTTGTGGAGACCCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCysLeuCysGlyLeuH 539
Db 963 CCTCCCTGCAGCAACCCACAGCTCTCAGAGACGACGAGAGGCCCTCTACTC-----CTGC 1016
QY 539 IsLeuGluSerSerGlnPheProPro 547
Db 1017 ACGCT-----CTTCCA 1027

RESULT 8

US-10-301-480-731556
; Sequence 731556, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 731556
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-731556

Alignment Scores:
Pred. No.: 2.05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 10 Gaps: 13

US-10-071-838-2 (1-549) x US-10-301-480-731556 (1-1052)

QY 277 lIeSerLeuGlYleuThrLeuArgLeuTrpAspValTyrLeuValGluglYgluGlnAla 296
Db 7 ATCTCTCTCGGGCTCACCCCTGCGCTGTGGACGTGTATCTGTGTAGAAAGCGAACAGCGC 66
QY 297 LeuMetProIleThrArgIleAlaPheIysValGlnGln----- 309
Db 67 TTGATGCCGATACAGAATCGCCTTTAAGTTCAAGCAGAGTAAGTCTACGTGTGCCAG 126
QY 309 ----- 309

Db 127 CGGGGCTGGGAGCCCTGGGGTCAGACCCCGACTGGCCCGAGGGCAGCTTCCTCACT 186
QY 309 -----
Db 187 GTCCTCATGATCCKCTGTCTTGCCAGAGGAGGTCTGGCCAGTGGGCTGGGAGAC 246
QY 309 -----
Db 247 ACTGTGACACCGAGCCCATCCCCACATGACCCAGATGAAAGTCGAGAGTGTGTAGCA 306
QY 309 -----
Db 307 CTTCCCTGTCCGATCGCCCCCAGCAGACTCTCTGTATATCTGACACCTGGGGT 366
QY 309 -----
Db 367 GGGCACAAGGATCCGGCACCAGCCAGTAGAGACTGAAGTGGCCAAGGATATAGCT 426
QY 310 -----Lys-ArgLeuTh 313
Db 427 GTGACCATTCACAGTAACCTCCCTGGCTGATATCCACCCCTGCTCAGAGCGCTCAC 486
QY 313 rlysthrserargcysglyprotrpalaargphecysasnargphevalaspthrtpral 333
Db 487 GAAGACGTCAGGTGTGGCCCGTGGGACGTTTTCGAACCGGTTGTTGATACCTGGGC 546
QY 333 aargaspgluaspthrvalleuylshleuargalasermetylsleuthrargly 353
Db 547 CAGGATGAGGACACTGTGCTCAAGCACTTAAGGCTCTATGAAGAACTAACAGAGAA 606
QY 353 sglnglyaspheuopropro-AlalysprogluglnlyserSerAlaserargprov 373
Db 607 GCAGGGGAGCTGCCACCCCGCAGGTG-----GGCTCCAGTGCCCATGTCCCT- 653
QY 373 alProAlaserarglyglylysthrleucyslysglyasparglnalapro-Progly 392
Db 654 --CCCATGTCACTCTGGGGTAGTCAGTAGAGGA-----GTGCCCCGGA 699
QY 393 ProProAlaargPheProargProiletrpserAlaserProProargAlaProargSer 412
Db 700 CCGGCAACCTACTACTGGGCT-----TCCTCTTCACCTTTCTTCT- 744
QY 413 SerthrProCysproglyglyalavalarggluaspthrtyrProvalglythrglnly 432
Db 745 -----CCTCTTCTCTCTGACT-CTAAGAAAGTACAGAGGCCCAACCGTCTCAGGGC 797
QY 433 ValProserProAlaleuAlaglnlyglyProglnglySerTrpArgPheLeuGlnTrp 452
Db 798 AGGCGCTCAGTGGCT-----GTATACTGG 821
QY 453 AsnSerMetProArgLeuProThrAspLeuAspValGlnlyProTrpPheArgHisTyr 472
Db 822 ACATGCTGTGCAAGC----- 836
QY 473 AspPheArgGlnSerCysTrpValargAlaIleSerGlnlyuaspglnleuAlaproCys 492
Db 837 -----AGGAGGGGATGTGGCAAG----- 857
QY 493 TrpGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db 858 -----CCCTCCAAAGAGCCCTCCACTTTCACGAGTGTCTCCCTCTCC 902
QY 513 SerAspGlnly-----ThrProPheArgAlaArgaspGlu 524
Db 903 CCTCGCAGGGCCCTCAAGTTACTAGACGAGCCAGACCATTTGTGGAGAGACCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCysleuCysglyleuH 539
Db 963 CCTCCCTGCAAGCACCACAGCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1016
QY 539 isleuGlnSerSerGlnPheProPro 547

Db 1017 ACGCT-----CCTCCA 1027
RESULT 9
US-10-301-480-731557
; Sequence 731557, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 731557
; LENGTH: 1052
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-731557
Alignment Scores:
Pred. No.: 2.05e-14 Length: 1052
Score: 380.00 Matches: 133
Percent Similarity: 34.7% Conservative: 10
Best Local Similarity: 32.3% Mismatches: 60
Query Match: 12.7% Indels: 210
DB: 10 Gaps: 13
US-10-071-838-2 (1-549) x US-10-301-480-731557 (1-1052)
QY 277 lleserleuglyleuthrleuargleuTrpaspValtyrleuvalgluglnAla 296
Db 7 ATCTCTCGGGCTCACCTCGGCTGTGGACGTGTATCTGTAGAGGCGAGAGCG 66
QY 297 leuMetProilethrargilealaphelysValGln----- 309
Db 67 TTGATGCCGATTAACAAGATCGCTTTAAGTTACAGAGTAAGTCTAGTGTGCCAG 126
QY 309 ----- 309
Db 127 CGGGGCTGGGAGCCCTGGGGTCAGACCCGACTGGCCCGAGGGCAGCTTCCTCACT 186
QY 309 ----- 309
Db 187 GTCCTCATGATCCTCTGTCTGGCCAGAGGAGGTCTGAGGCTGGGAGAGAC 246
QY 309 ----- 309
Db 247 ACTGTGACACCGAGCCCATCCCCACATGACCCAGATGAAGTCGAGAGTGTGTAGCA 306
QY 309 ----- 309
Db 307 CTTCCCTGTCCGATCGCCCCCAGCAGACTCTCTGTATATCTGAGACACTGGGGT 366
QY 309 ----- 309
Db 367 GGGCACAAGGATCCGGCACCAGCCAGTAGAGACTGAAGTGGCCACGGGATATAGACT 426
QY 310 -----Lys-ArgLeuTh 313
Db 427 GTGACCATTCACAGTAACCTCCCTGGCTGATATCCACCCCTGCTCAGAGCGCTCAC 486
QY 313 rlysthrserargcysglyprotrpalaargphecysasnargphevalaspthrtpral 333
Db 487 GAAGACGTCAGGTGTGGCCCGTGGGACGTTTTCGAACCGGTTGTTGATACCTGGGC 546
QY 333 aargaspgluaspthrvalleuylshleuargalasermetylsleuthrargly 353

Db 547 CAGGATGAGGACACTGTGCTCAAGCACTTAGGGCTCTATGAAAGAACTAACAAAGAAA 606
QY 353 BGLNGIYAAPLeuProProPro-AlaIysProGIuNGIYSerSerAlaSerArxProV 373
Db 607 GCAGGGGGAGCTGCGCAACCCCAAGGTG-----GGCTCCAGTGCCATGTCCCT- 653
QY 373 aIProAlaSerArxGIYGIYLYSThrLeuCybLYSGLYAAspArgGlnAlaPro-ProGIY 392
Db 654 --CCCATGTCAACCTCTGGGGTAGTCAGTAGTAGGGA-----GTGCCCGGGA 699
QY 393 ProProAlaArgPheProArgProIleTrpSerAlaSerProProAlaArgSer 412
Db 700 CCCGCAACCTACTACTGCGGCT-----TCCTCTTCACTTTTCTTCT- 744
QY 413 SerThrProCybProGIYGIYAlaValArgGluAspThrTYrProValGIYThrGIY 432
Db 745 -----CCTCTTCTCTCTGAGCT-CTAAGAAAGTACAGAGGCCCAACGGTCTCAGGGC 797
QY 433 ValProSerProAlaLeuAlaGlnGIYGIYProGIYSerTrpArgPheLeuGlnTrp 452
Db 798 AGCGGCTCAGTGGT-----GTATACCTGG 821
QY 453 AsnSerMetProArgLeuProThrAspLeuAspValGIYGIYProTrpPheArgHisTYr 472
Db 822 ACATGCTGTGCACGC----- 836
QY 473 AspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAlaProCys 492
Db 837 -----AGAGGGGGAGTGTGGCAAGA----- 857
QY 493 TrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThrAsp 512
Db 858 -----CCCTCCAAACAAGCCCCCTCCACTTTCACGGGTGTCTCCCTCTCC 902
QY 513 SerAspGlnGIY-----ThrProPheArgAlaArgAspGlu 524
Db 903 CCCTCGACAGGGCCCTCAAGTTACTAGACGAGCCACCATTTGTGGAGACCCCGCC 962
QY 525 GlnProCys--AlaProThrSer-----GlyProCybLeuCybGIYLeuH 539
Db 963 CCTCCCTGCAGACCAACCAAGCCTCAGAGAGCAGAGAGCCCTCACTC-----CTGC 1016
QY 539 IseLeuGlnSerGlnPheProPro 547
Db 1017 ACGCT-----CCTCCA 1027
RESULT 10
US-09-925-065A-55408
; Sequence 55408, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55408
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-925-065A-55408
Alignment Scores:
Pred. No.: 1.53e-13 Length: 1294
Score: 367.00 Matches: 143
Percent Similarity: 32.9% Conservative: 22
Best Local Similarity: 28.5% Mismatches: 86
Query Match: 12.3% Indels: 252
DB: 6 Gaps: 16
US-10-071-838-2 (1-549) x US-09-925-065A-55408 (1-1294)
QY 202 ProGIuGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeu 221
Db 34 CCAGCCTCATGGGCGAGACTGGGCCAGAGCCGACTTGAGAGGGCTCAG--GGAGCCTC 90
QY 222 GlnGlyPheHisSerProAsnGIYGIYThrValGlnGIYLeuGlnAspGlnGlnHis 241
Db 91 AAGCCTGGGCAAGCCCTCT-----CTCCAGAGCCACATCCCCAC 132
QY 242 -----ValValAlaThrSerGlnPro----- 248
Db 133 TCMAATGAGTGCTCCCATGAGAGCTTCAAGACCTTGTGTGACCCAGCGTCTGGAGGG 192
QY 249 -----LysThrMetGlnHisGlnAspLYSAspLeuCys 260
Db 193 CTCAGGCAACCTCATGGGAGAGAGGTCACTGACTCAGAGACTG-AAAGCCCAAGTGT 251
QY 261 GlnGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274
Db 252 GCGAGCTCGAGCCACCAAGCCCCCGCTGGAAGACCAAGTCTTTCACACCTGTGTCC 311
QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTYrLeuValGIYGIY 294
Db 312 CCACAGATCTCTCTCGGGCTCACCCCTGCGCTGTGGAGCGTATCTGTAGAAGGCGAA 371
QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLYSValGln-Gln----- 309
Db 372 CAGCGTGTGATGCCGATACCAAGATCGCCTTTAAGGTTCAGTAGAGTAAGTCTACGTGT 431
QY 309 ----- 309
Db 432 GCCCAGCGGGCCTGGGGAGCCCTGGGGTCAAGCCCGACTGGCCGAGGGCAGCTTCT 491
QY 309 ----- 309
Db 492 CACACTGTCTCATGATCTCTGTCTGGCCAGAGGAGGTCTGGCCAGGTGGGTGGG 551
QY 309 ----- 309
Db 552 CAGGACCTGTGACACCGAGCCCATYCCCCACATGACCCAGATGAAAGTCGAGAGTGTGG 611
QY 309 ----- 309
Db 612 TGAGCACTTCTCTGTCCAGATCACCCCCAGCCACAGTCTCTGTATATCTGACGCC 671
QY 309 ----- 309
Db 672 TGGGTGGCCAAAGATCCGGCACCGCCAGTGGGAGACTGAAGTGGCCAGGGGTA 731
QY 310 -----LysAr 311
Db 732 TGAGCTGTGACCAATCCAGGTAATCCCTGGCCTGATATCAACCTGTCCCTAGAGCG 791
QY 311 gLeuThrLYrThrSerArgCysGIYProTrpAlaArgPheCysAsnArgPheValAspTh 331
Db 792 CCTCAGGAAGACGTCCAGGTGTGGCCCGTGGGACGTTTGTGAACCGGTCTGTGATGC 851
QY 331 rTrpAlaArgAspGluAspThrValLeuLYSHisLeuArgAlaSerMetLYSLeuTh 351
Db 852 CTGGGCCAGGAGTATGACACTGTGTCAAGCATCTTAGGGCTCTATGAAGAACTAAC 911
QY 351 rArgLYSGlnGlyAspLeuProProPro-AlaIysProGIuNGIYSerSerAlaSerA 371

Db 912 AAGAAAGCAGGGGACCTGCCACCCCAAGGTG-----GGCTCCAGTGCATGT 959
QY 371 rgProValProAlaSerArgGlyGlyLeuThrLeuCysLysGlyAspArgGlnAlaPro- 390
Db 960 CCCCT---CCCATGTCAACCTCTGGGGTAGTCAATAGTAGGGGA-----GTGC 1004
QY 391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProAlaPro 410
Db 1005 CCGGACCCCGCAACCTACTACTGAGGCT-----TCCTCTTCACTTTCTTCTCT 1055
QY 411 ArgSerSerThrProCysProGlyGlyAlaValArgGlyAspThrTrpProValGlyThr 430
Db 1056 -----CCTCTTCTCTCTGAGCT-CTAAGAAAGTACAGAGGCCACCGGTCT 1102
QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450
Db 1103 CAGGCGAGCGCTGCTGCGCT-----GTG 1126
QY 451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470
Db 1127 TACTGACATGCTGTGCACGC----- 1147
QY 471 HisTrpAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGlyAspGlnLeuAla 490
Db 1148 -----AGGAGGGGATGTGGCGCAGA----- 1168
QY 491 ProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaAlaProSer 510
Db 1169 ----- 1174
QY 511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530
Db 1175 -----AACAGCCCCCT----- 1186
QY 531 SerGlyProCysLeuCysGlyLeuHisLeuGluSerSerGlnPheProProGlyPhe 549
Db 1187 -----CCCATTTCTTGGCGGTGTCTCTCTCCCTCCCGCAGGGCCCTCCAAGTTAC 1237

RESULT 11

US-09-925-065A-55409
; Sequence 55409, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55409
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-55409

Alignment Scores:

Pred. No.: 1.53e-13 Length: 1294
Score: 367.00 Matches: 143
Percent Similarity: 32.9% Conservative: 22
Best Local Similarity: 28.5% Mismatches: 86

Query Match: 12.3% Indels: 252
DB: 6 Gaps: 16
US-10-071-838-2 (1-549) x US-09-925-065A-55409 (1-1294)
QY 202 ProGlnGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeu 221
Db 34 CCAACCTCATGGGAGACTGGGCCAGACCCGACTTGAGAGGGCTCAG--GGAAGCCTC 90
QY 222 GlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnHis 241
Db 91 AAGCCCTGGGCAAGCCCTCT-----CTCAGGAGCCACATCCCCAC 132
QY 242 -----ValValAlaThrSerGlnPro----- 248
Db 133 TCAATGAGTGCCTCCCATGAGAGCTTCAAGACCTGTGTGACCCAGCCTCTGAGAGG 192
QY 249 -----LysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 193 CTCAGGCGACCTTCATGGGGAGGAGGAGTCACTGACTCAGAGACTG-AAAGCCAGTGT 251
QY 261 GlyGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274
Db 252 GCGCAGCTCGAGCCACAGCCCGCAGCTGGAAGACCAAGTTCCTTTCACACCTGTCTCC 311
QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTrpLeuValGlnGly 294
Db 312 CCACAGATCTGTCTGGGCTCACCCCTGGCCTGTGGACGTATCTGTGTAAGAAGCGAA 371
QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309
Db 372 CAGCGTGTGATGCGGATACAAGAATCGCTTAAAGTTCAGTAGAGTAAGTCTACGTGT 431
QY 309 ----- 309
Db 432 GCCCAGCGGGGCTGGGAGCCCTGGGGTCAAGCCCGACTGGCCGAGGGCAGCTTCT 491
QY 309 ----- 309
Db 492 CACACTGTCTCATGATCTCTGTCTGGCCAGAGGAGTCTGGCCAGGTGGCTGG 551
QY 309 ----- 309
Db 552 CAGACACTGTGACACCGAGCCCATTTCCACATGACCCAGATGAAGTCGAGAGTGTGG 611
QY 309 ----- 309
Db 612 TGAGCACTTCCCTGTCCAGATCRCCCCCAGCCACAGTCTCTGTATATCTGACGCC 671
QY 309 ----- 309
Db 672 TGGGTGGCCCAAAAGGATCCGGCAGCCCGCAGTGGAGACTGAAGTGGCCACGGGTA 731
QY 310 -----LysAr 311
Db 732 TGAGCTGTGACCATTTCCAGGTAATCCCTGGCTGTATATCCACCTGTCTCTAGAGCG 791
QY 311 GLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspTh 331
Db 792 CCTCACAGACAGTCCAGGTGTGCGCCGTGGCAGCTTTTGGAAACCGGTCTGTGATGC 851
QY 331 rTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuTh 351
Db 852 CTGGGCGCAGGATGATGACACTGTGTCTCAAGCATCTTAGGCCCTCTATGAAGAACTAAC 911
QY 351 rArgLysGlnGlyAspLeuProProPro-AlaLysProGlnGlnGlySerSerAlaSer 371
Db 912 AAGAAAGCAGGGGACCTGCCACCCCAAGGTG-----GGCTCCAGTGCATGT 959
QY 371 rgProValProAlaSerArgGlyGlyLeuThrLeuCysLysGlyAspArgGlnAlaPro- 390
Db 960 CCCCT---CCCATGTCAACCTCTGGGGTAGTCAATAGTAGGGGA-----GTGC 1004

QY 391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProAlaPro 410
Db 1005 CCGGACCCGCAACCTACTACTGGGCT-----TCCCTTTCACCTTTCTTCT 1055
QY 411 ArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThr 430
Db 1056 -----CCTCTTCTCTCTGACT-CTAAGAAAGTACAGAGGCCACCGGTCT 1102
QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450
Db 1103 CAGGCGAGCGCTCAGTGCCT-----GTA 1126
QY 451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470
Db 1127 TACTGACATGCTGTGCACGC----- 1147
QY 471 HisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAla 490
Db 1148 -----AGGAGGGGGATGTGGCGAGA----- 1168
QY 491 ProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaAlaProSer 510
Db 1169 -----CCCTCC 1174
QY 511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530
Db 1175 -----AACAGCCCCCT----- 1186
QY 531 SerGlyProCysLeuCysGlyLeuHisLeuGlnSerSerGlnPheProProGlyPhe 549
Db 1187 -----CCCACTTTCTGCGGTGTCTCCCTTCCCTCCGACGGGCCCTCCAGTTAC 1237

RESULT 12

US-10-301-480-156646
; Sequence 156646, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156646
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-156646

Alignment Scores:

Pred. No.: 1.53e-13 length: 1294
Score: 367.00 Matches: 143
Percent Similarity: 32.9% Conservative: 22
Best Local Similarity: 28.5% Mismatches: 86
Query Match: 12.3% Indels: 252
DB: 9 Gaps: 16

US-10-071-838-2 (1-549) x US-10-301-480-156646 (1-1294)

QY 202 ProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeu 221
Db 34 CCAGCCTCATGGGCAGACTGGGCCAGGACCCGACTTGAGAGGGCTCAG---GGAAGCCTC 90
QY 222 GlnGlyPheHisSerProAsnGlnGlyThrValGlnGlyLeuGlnAspGlnGlnHis 241
Db 91 AAGCCCTGGGCAAGCCCTCT-----CTCAGGAGCCACATCCCCAC 132

QY 242 -----ValValAlaThrSerGlnPro----- 248
Db 133 TCAAATGAGTGCTCCCATGAGAGCTTCAAGACCTTGTCTGACCAGCGTCTGGAGGG 192
QY 249 -----LysThrMetGlnHisGlnAspLysLysAspLeuCys 260
Db 193 CTCAGGCGACCTCATGGGAAGAGGTACTGACTCAGAGACTG-AAAGCCAGTGT 251
QY 261 GlyGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274
Db 252 GCGCAGCTCGAGCCACCAGCCCCAGCTGGAAGACCAAGTTCTTTCACACCTGTGTCC 311
QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlu 294
Db 312 CCACAGATCTTCTCGGGCTCACCTCGCCTGTGGAGCTGTATCTGTAGAAGCGGAA 371
QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309
Db 372 CAGCGTTGATGCCGATTAACAAAGATCGCTTTAAGTTCAAGTAAGTCTACGTGT 431
QY 309 ----- 309
Db 432 GCCCAGCGGGCCTGGGAGCCTGGGGGTACAGCCCGCACTGGCCGAGGCGAGTTCTCT 491
QY 309 ----- 309
Db 492 CACACTGTCTCATGATCTCTGTCTGGCCAGAGGAGGTCTGCCAGTGGCTGGG 551
QY 309 ----- 309
Db 552 CAGGACACTGACACCGAGCCCATYCCCAATGACCCAGATGAAGTCGAGAGTGTGG 611
QY 309 ----- 309
Db 612 TGAGCACTTCCCTGTCCAGATCACCCCCAGCCACAGTCTCTGTATATCTGACGCC 671
QY 309 ----- 309
QY 310 -----LysAsr 311
Db 732 TGAGCTGTGACCATTCCAGGTAACTCCCTGGCCTGATATCCACCTGTCCCTAGAGCG 791
QY 311 gLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspTh 331
Db 792 CCTCAGAAAGAGTCCAGGTGTGGCCGTGGGCACGTTTTCGAACCGGTTGTTGATGC 851
QY 331 rTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuTh 351
Db 852 CTGGGCCAGGGATGATGACACTGTGCTCAAGCATCTTAGGGCCTCTATGAAGAACTAAC 911
QY 351 rArgLysGlnGlyAspLeuProProPro-AlaLysProGlnGlnGlySerSerAlaSerA 371
Db 912 AAGAAAGCAGGGGAGCTGCCACCCCGAGGTG-----GGCTCCAGTGCATGT 959
QY 371 rgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPro- 390
Db 960 CCCCT--CCCATGTACACCTCTGGGGTAGTCAATAGTAGGGA-----GTGC 1004
QY 391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProAlaPro 410
Db 1005 CCGGACCCGCAACCTACTACTGGGCT-----TCCCTTTCACCTTTCTTCTCT 1055

QY 411 ArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThr 430
Db 1056 -----CCTCTTCTCTCTGACT-CTAAGAAAGTACAGAGGCCACCGGTCTCT 1102
QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450
Db 1103 CAGGCGAGCGCTCAGTGCCT-----GTA 1126
QY 451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470

```
Db      1127 TACTGGACATGCTGTGCACGC----- 1147
QY      471 HieTYrAspRheArgGlnSerCystrPrvaAlaAlaIleSerGlnIuAspGlnLeuAla 490
Db      1148 -----AGGAGGGGGATGTGGCGCAGA----- 1168
QY      491 ProCystrPrGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAlaProSer 510
Db      1169 ----- 1174
QY      511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530
Db      1175 -----AACAGCCCCCT----- 1186
QY      531 SerGlyProCysLeuCyGlyLeuHisLeuGlnSerSerGlnPheProGlyPhe 549
Db      1187 -----CCCACTTCTGCGGTGTCTCCCTCCCCCTCGCAGGGCCCTCCAAGTTAC 1237
```

RESULT 13
US-10-301-480-156647
; Sequence 156647, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156647
; LENGTH: 1294
; TYPE: DNA
; ORGANISM: Homo sapien
; US-10-301-480-156647

Alignment Scores:
Pred. No.: 1.53e-13 Length: 1294
Score: 367.00 Matches: 143
Percent Similarity: 32.9% Conservative: 22
Best Local Similarity: 28.5% Mismatches: 86
Query Match: 12.3% Indels: 252
DB: 9 Gaps: 16

```
US-10-071-838-2 (1-549) x US-10-301-480-156647 (1-1294)
QY      202 ProGlnIuAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnIuArgHisSerLeu 221
Db      34 CCAGCCTCATGGGCAGACTGGGCCAGGACCCGACTTGAGAGGGCTCAG---GGAAGCCTC 90
QY      222 GlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnIuHis 241
Db      91 AAGCCCTGGGCAAGCCCTCT-----CTCCAGGAGCCACATCCCCAC 132
QY      242 -----ValValAlaThrSerGlnPro----- 248
Db      133 TCMAATGATGCTCCCATGAGAGCTTCAAGACCTTGTGTGACCCAGCGTCTGGAGGG 192
QY      249 -----LysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db      193 CTCAGGCGACCTCATGGGGAAGAGGTCACTGACTCAGAGACTG-AAAGCCCACTGT 251
QY      261 GlyGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274
Db      252 GCGCAGCTCGAGCCACAGCCCAAGCCTGGAAGACCAAGTTCTTTCACACCTGCTGCC 311
QY      275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlu 294
```

```
Db      312 CCACAGATCTCTTCGGGCTCACCCCTGGCCCTGTGGAGCGTGTATCTGTAGAAGCGCA 371
QY      295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309
Db      372 CAGCGTTGATGCCGATTAACAAGAAATGCCCTTTAAGGTTCAGTAGAGTAAGTCTACGTGT 431
QY      309 ----- 309
Db      432 GCCCAGCGGGCCTGGGGAGCCCTGGGGTCAAGACCCGACTGGCCCGAGGCAGCTTCCT 491
QY      309 ----- 309
Db      492 CACACTGCTCATGATCCTCTGTCTGTGGCCCAAGAGGAGGTCTGGCCAGGTGGGCTGGG 551
QY      309 ----- 309
Db      552 CAGGACACTGTGACACCGAGCCCATTTCCCACATGACCCAGATGAAGTGAAGAGTTGG 611
QY      309 ----- 309
Db      612 TGAGCACTTCCTGTCCAGATCRCCCCCAGCCACAGTCTCCTGTATATCTGAGCGCC 671
QY      309 ----- 309
Db      672 TGGGTGGCCACAAGAAGATCCGGCACCGCCAGTGGAGACTGAAGTGCCACGGGTA 731
QY      310 -----LysAr 311
Db      732 TGAGCTGTGACCATTCACAGTACTCCCTGGCCTGATATCCACCCTGTCCCTAGACG 791
QY      311 GlnLeuThrLysThrSerArgCyGlyProTrpAlaArgPheCysAsnArgPheValAspTh 331
Db      792 CCTCACGAAGACGTCAGAGTGTGGCCCGGTGGCACGTTTGTGAACCGGTTCTGTATGC 851
QY      331 rTrpAlaArgAspGlnAspThrValIleuLysHisLeuArgAlaSerMetLysLysLeuTh 351
Db      852 CTGGGCCAGGATGATGACACTGTGCTCAAGCATCTTAGGGCCTTATGAAGAACTTAAC 911
QY      351 rArgLysGlnGlyAspLeuProProPro-AlaLysProGlnGlnGlySerSerAlaSerA 371
Db      912 AAGAAAGCAGGGGACCTGCCACCCCAAGTG-----GGCTCCAGTGCATGT 959
QY      371 rGProValProAlaSerArgGlyGlyLysThrLeuCyLysGlyAspArgGlnAlaPro- 390
Db      960 CCCCT---CCCATGTCAACCTCTGGGGTAGTCAATAGTAGGGA-----GTGC 1004
QY      391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPro 410
Db      1005 CCGGGACCCGCAACCTACTACTGGGCT-----TCCTCTTCACCTTTTCTTCCT 1055
QY      411 ArgSerSerThrProCysProGlyGlyAlaValArgGlnAspThrTyrProValGlyThr 430
Db      1056 -----CCTCTTCTCTGTAAGT-CTTAAGAAATGACAGAGGCCACCGGTCTCT 1102
QY      431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450
Db      1103 CAGGCGAGCGCTCAGTGGCT-----GTA 1126
QY      451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470
Db      1127 TACTGGACATGCTGTGCACGC----- 1147
QY      471 HieTYrAspRheArgGlnSerCystrPrvaAlaAlaIleSerGlnIuAspGlnLeuAla 490
Db      1148 -----AGAGGGGATGTGGCGCAGA----- 1168
QY      491 ProCystrPrGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAlaProSer 510
Db      1169 -----CCCTCC 1174
QY      511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530
```

Db 1175 -----ACAAAGCCCCCT----- 1186

QY 531 SerGlyProCybLeuCybGlyLeuHisLeuGluSerSerGlnPheProGlyPhe 549

Db 1187 -----CCCACTTTCTGCGGTGTCTCCCTCTCCCTCGCAGGGCCCTCCAAAGTTAC 1237

RESULT 14

US-10-301-480-770055

; Sequence 770055, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 770055

; LENGTH: 1294

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-770055

Alignment Scores:

Pred. No.:	1.53e-13	Length:	1294
Score:	367.00	Matches:	143
Percent Similarity:	32.9%	Conservative:	22
Best Local Similarity:	28.5%	Mismatches:	86
Query Match:	12.3%	Indels:	252
DB:	10	Gaps:	16

US-10-071-838-2 (1-549) x US-10-301-480-770055 (1-1294)

QY 202 ProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnArgHisSerLeu 221

Db 34 CCAGCCTCATGGGCACTGGGCGCAGACCCTGAGAGGCTGAG---GGAAGCCTC 90

QY 222 GlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlnHis 241

Db 91 AAGCCCTGGCAAGCCCTCT-----CTCCAGAGCCACATCCCCAC 132

QY 242 -----ValValAlaThrSerGlnPro----- 248

Db 133 TCAATAGAGTCCCTCCCATGAGAGCTTCAAGACCTTGCTGACCCAGCGTCTGAGGG 192

QY 249 -----LysThrMetClyHisGlnAspLysLysAspLeuCys 260

Db 193 CTCAGGCGACCTCATGGGGAAGGAGTCACTGACTCAGGAGACTG-AAAGCCCACTGT 251

QY 261 GlyGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274

Db 252 GCGCAGCTCGAGCCACCAAGCCCCAGCCTGGAAGGACCAAGTCTTTACACCTGCTGTCC 311

QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlu 294

Db 312 CCAAGATCTCTCGGGCTCAACCTGCGCTGTGGGACGTGTATCTGTGTAAGAGCGAA 371

QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309

Db 372 CAGGCGTGTATGCGGATACAAGAATCGCCTTAAGTTCAGTAGAGTAAGTCTACGTGT 431

QY 309 ----- 309

Db 432 GCCCAAGCGGGCCTGGGGAAGCCCTGGGGTCAAGACCCGACTGGCCGAGGAGCTTCCT 491

QY 309 ----- 309

Db 492 CACACTGTCTCATGATCCTCTGTCTGTGGCCAGAGGAGGTCTGGCCAGGTGGGCTGGG 551

QY 309 ----- 309

Db 552 CAGACACTGTGACACCGAGGCCATYCCCCACATGACCCAGATGAAGTGCAGAGTGTGG 611

QY 309 ----- 309

Db 612 TGAGCACTTCCCTGTCCAGATGACCCCCAGCCACAGTCTCTGTATATCTGACGCC 671

QY 309 ----- 309

Db 672 TGGGTGGCCAAAGAATCCGGCACCGCCAGTGGGAGACTGAAGTGGCCACGGGATA 731

QY 310 -----LysAr 311

Db 732 TGAGCTGTGACCATTTCCAGGTAACCTCCCTGGCCTGATATCCACCCTGTCCCTAGAGCG 791

QY 311 GLeuThrLysThrSerArgCybGlyProTrpAlaArgPheCysAsnArgPheValAspTh 331

Db 792 CCTCAGAGACGTCAGAGTGTGGCCGTGGGACGTTTGGAAACCGGTTCTTGATGC 851

QY 331 rTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuTh 351

Db 852 CTGGGCCAGGATGATGACACTGTCTCAAGCATCTTAGGGCTTATGAAGAACTAAC 911

QY 351 rArgLysGlnGlyAspLeuProProPro-AlaLysProGluGlnGlySerSerAlaSerA 371

Db 912 AAGAAAGCAGGGGACCTGCACCCCAAGTG-----GGCTCCAGTGCATGT 959

QY 371 rgProValProAlaSerArgGlyGlyLysThrLeuCybLysGlyAspArgGlnAlaPro- 390

Db 960 CCCCT---CCCATGTCAACCTCTGGGGTAGTCAATAGTAGGGA-----GTGC 1004

QY 391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPro 410

Db 1005 CCGGAGCCCGCAACCTACTACTGGGCT-----TCCTCTTCACTTTCTTCTCT 1055

QY 411 ArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThr 430

Db 1056 -----CCTCTTCTCTGTAAGTAAAGTACAGAGGCCACCGGTCTCT 1102

QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450

Db 1103 CAGGCGAGCGCTCAAGTGGT-----GTA 1126

QY 451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheArg 470

Db 1127 TACTGACATGCTGTGCACGC----- 1147

QY 471 HisTrpAspPheArgLysSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAla 490

Db 1148 -----AGGAGGGGATGTGGCGAGA----- 1168

QY 491 ProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaIleProSer 510

Db 1169 -----CCCTCC 1174

QY 511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaIleProThr 530

Db 1175 -----ACAAAGCCCCCT----- 1186

QY 531 SerGlyProCybLeuCybGlyLeuHisLeuGluSerSerGlnPheProGlyPhe 549

Db 1187 -----CCCACTTTCTGCGGTGTCTCCCTCTCCCTCGCAGGGCCCTCCAAAGTTAC 1237

RESULT 15

US-10-301-480-770056

; Sequence 770056, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; TITLE OF INVENTION: in the Human Genome
 ; FILE REFERENCE: 108827.137
 ; CURRENT APPLICATION NUMBER: US/10/301,480
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/215,598
 ; PRIOR FILING DATE: 2002-08-09
 ; PRIOR APPLICATION NUMBER: US 60/311,695
 ; PRIOR FILING DATE: 2001-08-10
 ; NUMBER OF SEQ ID NOS: 1226818
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 770056
 ; LENGTH: 1294
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; US-10-301-480-770056

Alignment Scores:

Pred. No.:	1.53e-13	Length:	1294
Score:	367.00	Matches:	143
Percent Similarity:	32.9%	Conservative:	22
Best Local Similarity:	28.5%	Mismatches:	86
Query Match:	12.3%	Indels:	252
DB:	10	Gaps:	16

US-10-071-838-2 (1-549) x US-10-301-480-770056 (1-1294)

QY 202 ProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeu 221
 Db 34 CCAGCCTCATGGGCGACACTGGGCGAGGAGCCGACTTGAGAGGGCTCAG--GGAAGCCTC 90
 QY 222 GlnGlyPheHisSerProAlaGlnGlyThrValGlnGlyLeuGlnAspGlnGlnGlnHis 241
 Db 91 AAGCCCTGGGCGAAGCCCTCT--CTCCAGAGCCACATCCCCAC 132
 QY 242 -----ValValAlaThrSerGlnPro----- 248
 Db 133 TCAATGATGCTCTCCATGAGAGCTTCAAGACCTGTGTGACCCAGCTCCTGAGGG 192
 QY 249 -----LysThrMetGlyHisGlnAspLysLysAspLeuGly 260
 Db 193 CTCAGGCGACCTCATGGGGAAGAGTCACTGACTCAGAGACTG-AAAGCCCACTGT 251
 QY 261 GlnGlnCysSerProLeuGly-----CysLeuIleArgIleLeuIle 274
 Db 252 GCGAGCTCGAGCCACAGCCCGCTGGAAGACCAAGGTTCTTTCACACCTGCTGTC 311
 QY 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGln 294
 Db 312 CCACAGATCTCTCTGGGCTCACCTGCGCTGTGGAGCTGTATCTGTAGAAAGCGCA 371
 QY 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGln-Gln----- 309
 Db 372 CAGCGTGTGATGCCGATTAACAAGATCGCTTTAAGGTTCAAGTAAGTCTACGTGT 431
 QY 309 ----- 309
 Db 432 GCCCAGCGGGGCTGGGAGCCCTGGGGTCAAGCCCGACTGGCCCGAGGAGCTTCT 491
 QY 309 ----- 309
 Db 492 CACACTGCTCATGATCCTCTGTTCGGCCAGAGGAGGTCTGGCCAGGTGGGCTGG 551
 QY 309 ----- 309
 Db 552 CAGGACACTGTGACACCGAGCCCATTTCCCATATGACCCAGATGAAGTGAAGTGTGG 611
 QY 309 ----- 309
 Db 612 TGAGCACTTCCTGTCCAGATCRCCCCCAGCCACAGTCTCTGTATATGTGACGCC 671
 QY 309 ----- 309
 Db 672 TGGGTGGCCACAAAAGATCCGGGACCGGCCAGTGGAGACTGAAGTGCCACGGGGTA 731

QY 310 -----LysArg 311
 Db 732 TGAGCTGTGACCATTTCCAGGTAACCTCCCTGCGCTGATATCCACCCTGTCCCTAGAGCG 791
 QY 311 GlnThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThr 331
 Db 792 CCTCAGAAAGACGTCCAGGTGTGCGCCGTGGGACAGTTTGTGAACCGGTTCTGTGATGC 851
 QY 331 rTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuThr 351
 Db 852 CTGGGCCAGGAGTATGACACTGTGCTCAACCATCTTAGGGCTCTATGAAGAACTAAC 911
 QY 351 rArgLysGlnGlyAspLeuProProPro-AlaLysProGluGlnGlySerSerAlaSer 371
 Db 912 AAGAAAGCAGGGGACCTGCCACCCCGAGTG-----GGCTCCAGTGCCATGT 959
 QY 371 rgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPro- 390
 Db 960 CCCCT--CCCATGTCAACCTCTGGGTAGTCAATAGTAGGGA-----GTGC 1004
 QY 391 ProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPro 410
 Db 1005 CCGGACCCGCAACCTACTACTGGGCT-----TCCCTTCACTTCTTCTCT 1055
 QY 411 ArgSerSerThrProCysProGlyValAlaValArgGluAspThrTrpProValGlyThr 430
 Db 1056 -----CCTCTTCTCTGAGCT-CTAAGAAAGTACAGAGGCCACCGGTCT 1102
 QY 431 GlnGlyValProSerProAlaLeuAlaGlnGlyProGlnGlySerTrpArgPheLeu 450
 Db 1103 CAGGCGAGGCGCTCAGTGCCT-----GTA 1126
 QY 451 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheArg 470
 Db 1127 TACTGACATGCTGTGCACGC----- 1147
 QY 471 HisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAla 490
 Db 1148 -----AGAGGGGAGATGTGGCGAGA----- 1168
 QY 491 ProCysTrpGlnAlaGlnHisProAlaGluArgValArgSerAlaPheAlaAlaProSer 510
 Db 1169 -----CCCTCC 1174
 QY 511 ThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnGlnProCysAlaProThr 530
 Db 1175 -----AACAAAGCCCT----- 1186
 QY 531 SerGlyProCysLeuCysGlyLeuHisLeuGluSerSerGlnPheProProGlyPhe 549
 Db 1187 -----CCCACTTCTGTGGGTGTCTCCCTCCCGCGAGGCGCTCCAAAGTTAC 1237

Search completed: April 5, 2006, 15:11:30
 Job time : 554 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:54:53 ; Search time 1234 Seconds
(without alignments)
3679.002 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVVEVAGSWMAQEREDITM.....TSGPCLGHLLESSQPPPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 segs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-O=/absas/ABSSWEB.spool/US10071838/runat_05042006_141920_5288/app_query.fasta_1
-DB=Published_Applications_NA_Main -QMT=fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs807
-USER=US10071838 @CGN_1_1_1026 @runat_05042006_141920_5288 -NCPU=6 -ICPU=3
-NO MAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2983	100.0	1964	5	US-10-071-838-1	Sequence 1, Appl1
2	2978	99.8	1993	8	US-10-839-882-31	Sequence 31, Appl1
3	2907.5	97.5	2146	5	US-10-071-838-3	Sequence 3, Appl1
4	2850	95.5	2084	7	US-10-302-172-692	Sequence 692, App
5	2827	94.8	2072	7	US-10-302-172-691	Sequence 691, App
6	2827	94.8	2647	6	US-10-108-260A-676	Sequence 676, App
7	2784	93.3	1752	6	US-10-094-466-61	Sequence 61, Appl

8	2773	93.0	1862	5	US-10-071-838-5	Sequence 5, Appl1
9	2191.5	73.5	7856	9	US-10-450-763-21335	Sequence 21335, A
10	2191.5	73.5	7856	9	US-10-450-763-24017	Sequence 24017, A
11	2191.5	73.5	7878	3	US-09-962-436-562	Sequence 562, App
12	2191.5	73.5	7878	9	US-10-843-641A-3021	Sequence 3021, App
13	2191.5	73.5	7878	9	US-10-956-157-1259	Sequence 1259, App
14	2133.5	71.5	8180	9	US-10-450-763-24016	Sequence 24016, A
15	2133.5	71.5	8201	7	US-10-283-975A-272	Sequence 272, App
16	2133.5	71.5	8201	8	US-10-278-698-640	Sequence 640, App
17	2133.5	71.5	8201	8	US-10-278-698-1037	Sequence 1037, App
18	2133.5	71.5	8284	9	US-10-450-763-21339	Sequence 21339, A
19	2133.5	71.5	8408	5	US-10-098-841-51	Sequence 51, Appl1
20	2133.5	71.5	8420	5	US-10-098-841-50	Sequence 50, Appl1
21	1795.5	60.2	3396	9	US-10-450-763-21338	Sequence 21338, A
22	1795.5	60.2	3904	9	US-10-450-763-24018	Sequence 24018, A
23	1732	58.1	2838	9	US-10-450-763-19813	Sequence 19813, A
24	1732	58.1	2838	9	US-10-450-763-22833	Sequence 22833, A
25	1732	58.1	2838	9	US-10-450-763-23326	Sequence 23326, A
26	1533.5	51.4	2424	9	US-10-450-763-23327	Sequence 23327, A
27	1507.5	50.5	9805	3	US-09-764-869-1824	Sequence 1824, App
28	1507.5	50.5	9805	3	US-09-764-891-7857	Sequence 7857, App
29	1507.5	50.5	9805	5	US-10-091-504-1824	Sequence 1824, App
30	1507.5	50.5	9805	6	US-10-227-577-1824	Sequence 1826, App
31	1500.5	50.3	9404	3	US-09-764-869-1826	Sequence 1826, App
32	1500.5	50.3	9404	3	US-09-764-891-7859	Sequence 7859, App
33	1500.5	50.3	9404	5	US-10-091-504-1826	Sequence 1826, App
34	1500.5	50.3	9404	6	US-10-227-577-1826	Sequence 1826, App
35	1469	49.2	9804	3	US-09-764-869-1829	Sequence 1829, App
36	1469	49.2	9804	3	US-09-764-891-7861	Sequence 7861, App
37	1469	49.2	9804	5	US-10-091-504-1829	Sequence 1829, App
38	1469	49.2	9804	6	US-10-227-577-1829	Sequence 1829, App
39	1465	49.1	9796	3	US-09-764-869-1825	Sequence 1825, App
40	1465	49.1	9796	3	US-09-764-891-7858	Sequence 7858, App
41	1465	49.1	9796	5	US-10-091-504-1825	Sequence 1825, App
42	1465	49.1	9796	6	US-10-227-577-1825	Sequence 1825, App
43	1428.5	47.9	6696	3	US-09-764-869-1821	Sequence 1821, App
44	1428.5	47.9	6696	3	US-09-764-891-7856	Sequence 7856, App
45	1428.5	47.9	6696	5	US-10-091-504-1821	Sequence 1821, App

ALIGNMENTS

RESULT 1
US-10-071-838-1
; Sequence 1, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071, 838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267, 615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1650)
US-10-071-838-1
Alignment Scores: 2.19e-265 Length: 1964
Pred. No.: 2.19e-265 Length: 1964

Score:	2983.00	Matches:	549
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	5	Gaps:	0

US-10-071-838-2 (1-549) x US-10-071-838-1 (1-1964)

QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
DB	1	ATGAGACGTGTAGAGGTCGCGGGCAGTTGTGTGGCAACAAGACGAGAGCATCATATTATG	60
QY	21	LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
DB	61	AAATACGAAGAAGGACACCCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCCTTTTGA	120
QY	41	SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu	60
DB	121	AGCTACAAACAACAGCTGCATCTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG	180
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
DB	181	ACTGCGCGGGAGCGCAAGCAAAATTCGGCGGGAGATCAGCCGGAAGACCAAGTGGGTGAT	240
QY	81	MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
DB	241	ATGCTGGGAGACTGGAGAAATACAAAGCAGACAGAAAGCTCATAGATCGAGCTACAG	300
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlnMet	120
DB	301	GGATGCCCCATGAACATCCGGGGCCGATGTGTGTCAGTCCCTCTGAACATTGAGAAATG	360
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
DB	361	AAGTTGAAAAACCCCGGAGATACCAATCATGAAGAGAGGCAAGAGTCATCTGAG	420
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
DB	421	CACATCCAGCCCATCGACCGGGAAGTAAGCGGGAATTAAAGAAACAATATTCTTCAGG	480
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
DB	481	GATCGATACGGAAACCAAGCAGCGGAACTACTCCACATCCTCTGGCAATATGAGGAGTAT	540
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr	200
DB	541	AAACCCGAGGTGGGCTACTGCAGGAGCTGAGCCACATCGCCGCTTGTCTCTCTAT	600
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
DB	601	CTTCCTGAGGAGATGCATCTCGGCACTGGTGACGTGCGCACTGAGAGGCACTCC	660
QY	221	LeuGlnGlyPheHisSerProAsnGlyLysIleThrValGlnGlyLeuGlnAspGlnGln	240
DB	661	CTGCAGGGAATTCACAGCCCAATGGCGGGACCGTCCAGGGGGCTCCAAGACCAACAGAG	720
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
DB	721	CATGTGTAGCAAGTCACAACCCCAACCATGGGGCATCAGACCAAGAAAGATCTATGT	780
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
DB	781	GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCTCGGG	840
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle	300
DB	841	CTCACCCCTGGCCTGTGGGAGCTGTATCTGGTAGAAGCGAACAAGGCTTGATGCCGATA	900
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
DB	901	ACAAAGAAATCGCTTTAAGTTCAAGCAAGACCGCCTCACGAAGACGTCCAAGTGTGGCCCG	960
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340

D	b		961	TGGGCA	C	GTTT	TTCACACC	GTTTCGTTGA	TAACTGGGCCAGG	GATGAGACA	CTGTGCTC	1020
O	y		341	LysHis	LeuArgAlaSerMet	LysIleuThrArg	LysGlnGlyAsp	LeuProPro	Pro			360
D	b		1021	AAGCAT	CTTAGGCCTCTATGA	GAACATAACA	GAAGAAGCAGGGG	ACTCGCACCCCA				1080
O	y		361	AlaLys	ProGlnGlnGlySer	SerAlaSerArg	ProValProAla	SerArgGlyLys				380
D	b		1081	GCCAAA	CCCGAGCAAGG	GTCGTCGGCAT	CCAGGCTGTG	CCGGCTTCA	CGTGCGGGAAG			1140
O	y		381	ThrLeu	CysLysGlyAsp	ArgGlnAlaPro	ProGlyProPro	AlaArgPhePro	ArgPro			400
D	b		1141	ACCCTT	GCAGAGGGGACAGG	CAAGCCCCCTTCA	GGCCACACAGCCC	CGGTTCCGCGGCC				1200
O	y		401	IleTrp	SerAlaSerPro	ProArgAlaPro	ArgSerSerThr	ProCysPro	GlyAla			420
D	b		1201	ATTGGT	CAAGCTTCCCGCAC	GGCACTCGTCTT	CCACACACCTGT	CCTGGTGGGGCT				1260
O	y		421	ValArg	GluAspThrTyrr	ProValGlyThr	GlnGlyValPro	SerProAla	LeuAlaGln			440
D	b		1261	GTCCGG	GAAGACACCTA	ACCCTGTGGCACT	CAAGGTGTGCCA	AGCCCGGCTTGCTCAG				1320
O	y		441	GlyGly	ProGlnGlySer	TrpArgPhe	LeuGlnTrp	PanSerMetPro	ArgLeuPro	Thr		460
D	b		1321	GGAGGA	CTCTCAGG	STTCTTGAGATT	CTTGAGTA	CTGGAATCCAT	GCCCCCTCCCAACG			1380
O	y		461	AspLeu	AspValGluGly	ProTrpPhe	ArgHisTyrAsp	PheArgGln	SerCysTrp	Val		480
D	b		1381	GACCTG	AGCGTAGAGGG	CCCTTGTTCCGCCA	TTATGATT	TCAGACAGAC	GTGCTGGGTC			1440
O	y		481	ArgAla	IleSerGlnGlu	AspGlnLeuAla	ProCysTrp	GlnAlaGln	HisProAla	Glu		500
D	b		1441	CGTGCA	TATCCACGAGG	AGACCACTGG	CCCCCTGCTGC	AGGCTGAACA	CCCTGCGGAG			1500
O	y		501	ArgVal	ArgSerAlaPhe	AlaAlaPro	SerThra	SerAspGln	GlyThrPro	PheArg		520
D	b		1501	CGGTGA	GATCGGCTTTCGCT	GCACCA	CCGACATGATT	CCGACCA	GGGCACCCCTT	CAGA		1560
O	y		521	AlaArg	AspGlnGlnPro	CysAlaPro	ThrSerGlyPro	CysLeuCys	GlyLeuHis	Ileu		540
D	b		1561	GCTAGG	AGACGAACAG	CCGTGTGCTCC	CACTCAGGG	CCCTTGCCCT	CTGCGGCTCCACT	TG		1620
O	y		541	GluSer	SerGlnPhePro	ProGlyPhe						549
D	b		1621	GAAGTT	CTTCAGTTCCCT	CCAGGCTTC	1647					

```

RESULT 2
US-10-839-882-31
; Sequence 31, Application US/10839882
; Publication No. US20040203106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: CORLEY, Neil C.
; APPLICANT: LAL, Preeti
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: JUNMING, Yang
; APPLICANT: SHIH, Leo L.
; TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
; FILE REFERENCE: PF-0619 PCT
; CURRENT APPLICATION NUMBER: US/10/839,882
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US/09/807,452
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/175,737; unassigned; 60/118,559; 09/249,740; unassigned
; 60/154,336
; PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11;
; 1999-04-22
;

```

; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 31
; LENGTH: 1993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4325626CB1
US-10-839-882-31

Alignment Scores:
Pred. No.: 6.45e-265 Length: 1993
Score: 2978.00 Matches: 548
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.8% Indels: 0
DB: 8 Gaps: 0

US-10-071-838-2 (1-549) x US-10-839-882-31 (1-1993)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 42 ATGACGTGGTAGAGGTCCGGGCACTTGGTGCGACAAAGACGAGACATCATTTATG 101
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 102 AAATACGAAGAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA 161
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 162 AGCTACAAACAACAGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCTCCTCTG 221
QY 61 ThrIleArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 222 ACTGCCGGGAGCGGAGCAAAATTCGGCGGAGATCACCCGAAAGACCAAGTGGGTGAT 281
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 282 ATGCTGGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCTTACAAG 341
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValleuLeuAsnIleGlnGluMet 120
DB 342 GGAATGCCCATGAACATCCGGGGCCGATGTGTGTCAGTCTCCTGAACACTGAGGAATG 401
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGlu 140
DB 402 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAAGAGGACAGAGTCATCTGAG 461
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 462 CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAAGAACATATATTCTTCAGG 521
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
DB 522 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCGCATATGAGAGTAT 581
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
DB 582 AACCCGAGGTGGGTACTGACGAGGACCTGAGCCACATCGCGCTTGTCTCTCTAT 641
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 642 CTTCCTGAGGAGGATGATCTGGGCACTGTGCACTGCTGGCCAGTGAGGCACTCC 701
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 702 CTGCAGGGATTTCACAGCCCAATGGCGGACCCGTCCAGGGGCTCCAAGACCAAGAG 761
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 762 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGACAAAGAAATCTATGT 821

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 822 GGGCAGTGTCCCCGTTAGGCTGCCTCATCCGGATATTGATTGACGGGATCTCTCGGG 881
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 882 CTCACCTCGCCCTGTGGGACGTGTATCTGTAGAACGGCAACAGCGCTTGATGCCGATA 941
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 942 ACAAGATTCGCTTTAAGGTTCAAGAGAGCGCTCAAGAAAGCTCCAGGTGTGGCCG 1001
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1002 TGGGCACTTTTTCAAACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1061
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1062 AAGCATCTTAGGGCTCTATGAAGAAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 1121
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1122 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCACTGGCGGGAAG 1181
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1182 ACCCTCTGCAAGGGGACAGGAGGCCCTCTCAGGCCCAACAGCCCGGTTCCGGCGGCC 1241
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1242 ATTGTGTACGTTCCCGGCCAGCGGACCTCGTTTTCACACCCCTGTCTGGTGGGCT 1301
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1302 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGCTGTGCCAGCCCGGCTGGCTCAG 1361
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1362 GAGAGACCTCAGGGTCTCTGAGAGATTCTGCAGTGGAATCCATGCCCGCTCCCAACG 1421
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1422 GACCTGAGCTAGAGGGCCTTGTTCCGCCATTATGATTTCAAGACAGCTGTGGTGC 1481
QY 481 ArgAlaIleSerGlnLysAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
DB 1482 CGTGCATATCCAGAGGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1541
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1542 CGGTTGAGATCGGCTTGTGCTGACACCAAGCACTGATTCGACCAAGGCAACCCCTCAGA 1601
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1602 GCTAAGGAGCAACAGCCGTGTCTCCACCTCAGGGCCTTGCTGTGGGCTCCACTTG 1661
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1662 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1688

RESULT 3
US-10-071-838-3
; Sequence 3, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838

; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2146
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1832)
US-10-071-838-3

Alignment Scores:
Pred. No.: 2.31e-258 Length: 2146
Score: 2907.50 Matches: 547
Percent Similarity: 89.7% Conservative: 0
Best Local Similarity: 89.7% Mismatches: 2
Query Match: 97.5% Indels: 62
DB: 5 Gaps: 1

US-10-071-838-2 (1-549) x US-10-071-838-3 (1-2146)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 1 ATGACGCTGTAGAGGTCCGCGGCGAGTTGGTGGCCAAAGACGAGACATCATTTATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 61 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCTTAAGCTTTTTCGA 120
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHis----- 53
DB 121 AGCTACAAACAACGTCGATCATTTGGGATTGTACAGTCCTCGCTCTGGAGTCA 180
QY 53 ----- 53
DB 181 GCCCCACAGAAAGCCCTGTCTCCCTTCCCTGTGCTTCTCTGGCTGAGCCCTGAG 240
QY 53 ----- 53
DB 241 CTGAAAGGGAACAGACGAGTCTCTTCTGGGGGTCCGCCACCGCTGGCGCTCCAG 300
QY 54 -----GluThrGluLeuProPro 59
DB 301 GCCCCGTGCACTTCTCAGCTCTGCTGGGTTGCCCTTACAGTGAAGACGGAGCTGCTCT 360
QY 60 LeuThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpVal 79
DB 361 CTGACTGCGGGAGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGT- 419
QY 80 AspMetLeuGlyAspTrpGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyr 99
DB 420 GATATGCTGGGAGACTGGGAGAAATCAAAAGCAGACAGAAAGCTCATAGATCAGGCTAC 479
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGlu 119
DB 480 AAGGGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAA 539
QY 120 MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysGlyLysArgSerSer 139
DB 540 ATGAAGTTGAAAAACCCCGGAAGATACCATGATCATGAAGGAAAGGGCAAGGTACT 599
QY 140 GlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 159
DB 600 GAGCACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAAGGAAGCATATATTCTTC 659
QY 160 ArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGlu 179
DB 660 AGGATCGATACGGAACCAAGCAGCGGGAATACTCCACATCTCTCTGGCATATGAGAG 719

QY 180 TyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeu 199
DB 720 TATAACCCGAGAGGTGGCTACTGCAGGACCTGAGCCACATCGCCGCTTGTCTCTC 779
QY 200 TyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 219
DB 780 TATCTTCTGAGGAGGATCATTTCTGGCAGCTGTGACGTGCTGGCCAGTGAAGGCAC 839
QY 220 SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 239
DB 840 TCCCTGCAGGATTTCACAGCCCAATGGCGGACCTCCAGGGGCTCCAAGACCAACAG 899
QY 240 GlnHisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeu 259
DB 900 GAGCATGTGTAGCCACGTCACAACCCAAAGCATGGGGCATCAGAGACAAAGAAATCTA 959
QY 260 CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu 279
DB 960 TGTGGCAGTGTCCCGCTTAGCTGCTCATCCGATATTGATTGACGGGATCTCTCTC 1019
QY 280 GlyLeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetPro 299
DB 1020 GGGCTACCCCTGCGCGCTGGGACGTGTATCTGTGAGAAGCGCAACAGCGGTGATGCCG 1079
QY 300 IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGly 319
DB 1080 ATAACAAGATCGCTTAAAGTTACAGACAGAGCGCTCACAGAACGTTCCAGGTGTGGC 1139
QY 320 ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 339
DB 1140 CCGTGGCAGCTTTTTCACCGGTTCTGTGATACCTGGGCCAGGAGTGAACAACACTGTG 1199
QY 340 LeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 359
DB 1200 CTCAAGCATCTTAGGGCTCTATGAAGAACTAAACAAGAACAGGGGACTGCCACCC 1259
QY 360 ProAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGly 379
DB 1260 CCAGCCAACCCGAGCAAGGTCGTCCGCATCCAGCCTGTGCCGCTTCACTGGCGGG 1319
QY 380 LysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg 399
DB 1320 AAGACCTCTGCAAGGGGGACAGGACCCCTCCAGGCCACAGCCCGCTTCCCGCGG 1379
QY 400 ProIleTrpSerAlaSerProProAlaProArgSerSerThrProCysProGlyGly 419
DB 1380 CCCATTGTGTCACTTCCCGGCCACGGGCACTGTTCTTCCACACCTGTCTGTGGG 1439
QY 420 AlaValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAla 439
DB 1440 GCTGTCCGGGAAGACACCTTACCTGTGGCAGTCAAGGTGTGCCAGCCCGCTGGCT 1499
QY 440 GlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuPro 459
DB 1500 CAGGAGGACCTCAGGCTTCTGGAAGATCTCGCAGTGAACCTCATGCCCCGCTCCCA 1559
QY 460 ThrAspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrp 479
DB 1560 ACGGACCTGGACGTAGAGGGCTTGTGTCCGCCATTATGATTTCAGACAGAGCTGCTGG 1619
QY 480 ValArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAla 499
DB 1620 GTCCGTGCATATCCACAGAGGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTTGG 1679
QY 500 GlnArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPhe 519
DB 1680 GAGCGGTGAGATCGGCTTTCGTGACCCAGCAGTGAATCCGACAGGGAACCCCTTTC 1739
QY 520 ArgAlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHis 539
DB 1740 AGAGCTAGGGAAGAACAGCCGTGTGCTCCACCTCAGGGGCTTGTGCTTGGCGCTCCAC 1799

Db 1609 CGGTGAGATCGCTTTCGTGCACCCAGCACTGATTCGACCAGGGCACCCTTCAGA 1668
QY 521 AIAARGAPGUGLGNPRGYSALAPRQTHRSERGLYPRQCYBLEUCYSGLYLEUHI SLEU 540
Db 1669 GCTAGGAGACGACGCGGTGTGCTCCCACTCAGGCGCTTGCTCTGCGGCTCCACTTG 1728
QY 541 GUSERSERGLNPHPRPROGLYPHE 549
Db 1729 GAAAGTCTCAGTTCCTCCAGGCTTC 1755

RESULT 5
US-10-302-172-691
; Sequence 691, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1e1 Arginine-rich Protein-like Nucleic Acids an
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 691
; LENGTH: 2072
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163) .. (1743)
US-10-302-172-691

Alignment Scores:
Pred. No.: 6.08e-251 Length: 2072
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 95.6% Mismatches: 2
Query Match: 94.8% Indels: 22
DB: 7 Gaps: 1

US-10-071-838-2 (1-549) x US-10-302-172-691 (1-2072)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGAGCGTGTAGAGGTTCGGGGCAGTTGTGGGCACAAGAGCGAGGACATCATTTATG 222
QY 21 LYSTYRGULYBSGLYHISARGALAGLYLEUPROGLUASPLYSGLYPROLYSPROPHARG 40
Db 223 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTGA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACACACACACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCTCCTCG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGGCGCAACCAATTCCGGCGGAGATCAGCCGAAAGACAAAGTGGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGAGAAATACAAAAGCAGCAAAAGCTCATAGATCGAGCGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCAATCCTCTGAACAACCTGAGAAATG 522

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTGAAAAAACC CGGAAGATACAGATCATGAAGAGAAAGGCAGAGGTCACTCTGAG 582
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 583 CACATCCAGCGCATCGACCGGGACGTAAAGCGGACATTAAGAAACATATATTCTTCAGG 642
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 643 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGCGCATATGAGAGTAT 702
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
Db 703 AACCCGAGGTGGCTACTGCAGGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 763 CTTCCTGAGAGATGATGATTCGGGCACTGGTGACGCTGCGCAGTAGAGCACTCC 822
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 823 CTGCAAGGATTTCAACAGCCCAATGGCGGAGACCGTCCAGGGGCTCCAAGACCAACAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 883 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCGGG 936
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 937 CTCACCTCGCGCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGCGTTGATGCGGATA 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACAAGATGCGCTTTAAGGTTCAAGAGCGCTCAAGAAAGCGTCCAGGTGTGCGCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1057 TGGGACGTTTTTTGCCAACCGGTTCTGTATACCTGGGCCAGGATGAGACACATGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1117 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGAAAGGAGGACCTGCACCCCCA 1176
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1177 GCCAAACCGAGCAAGGGTGTGCGCATCCAGGCTGTGCGGCTTCAAGTGGCGGAAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTCGCAAGGGGAGACAGGACGCGCTCCAGGCCCAACCCCGTTCCCGCGGCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1297 ATTTGGTCAGCTTCCCGGCCACGGGCACTCGTCTTCCACACACCTGTCTGTGGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1357 GTCCGGAGAGACACTTACCTGTGGGCACTCAGGGGTGTGCCACCGCGCTGTGCTCAG 1416
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1417 GGAGACCTCAGGGTCTCTGAGATTCTGTGAGTCCATGGAATCCATGCCCGCTCCCAAG 1476
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1477 GACCTGAGCTAGAGGGCCCTGTGTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 1536

[illegible]

```

RESULT 6
US-10-108-260A-676
; Sequence 676, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-676

```

Alignment Scores:	
Pred. No.:	8, 17e-251
Score:	2627.00
Percent Similarity:	95.6%
Best Local Similarity:	95.6%
Query Match:	94.8%
DB:	6
	Gaps: 1
	Length: 2647
	Matches: 525
	Conservative: 0
	Mismatches: 2
	Indels: 22

US-10-071-838-2 (1-549) x US-10-108-260A-676 (1-2647)

QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
Db	755	ATGACGTGTAGAGGTCCGGGCAGTTGTGTGGCACAAGACGAGACATCATTAAG	814
QY	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
Db	815	AAATACGAAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGGCTTAAGCCTTTCCGA	874
QY	41	SerTyrAsnAsnAsnValAspHisIleGluYTIleValHisGluThrGluLeuProProLeu	60
Db	875	AGCTACACACACACGTCGATCATTTGGGGATTGTACATGAGACCGAGCTGCTCTCTG	934
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
Db	935	ACTGCGGGAGCGCAACAAATTCCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT	994
QY	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
Db	995	ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAAAAGCTCATAGATCGAGCGTACAG	1054
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet	120
Db	1055	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACACTGAGGAATG	1114
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
Db	1115	AAATTGAAAAACCCCGAAGATACCAAGATCATGAAGGAGAAGGGCCAAGAGTCACTGAG	1174
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160

Db	1175	CACATCCAGCGCATGCACCGGACGTAAAGCGGGACATTAAAGAACATATATCTTCAGG	1234
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr	180
Db	1235	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCGCATATGAGAGTAT	1294
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr	200
Db	1295	AAACCGGAGGTGGGCTACTGCAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT	1354
QY	201	LeuProGluGluAspAlaPheThrPalaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	1355	CTTCCTGAGAGATGCATCTCTGGGCACCTGGTGCAGCTGCTGGCCAGTAGAGGCACTCC	1414
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	1415	CTGCAGGATTTCAACGCCCAATAGCGGAGCCGTCACAGGGGCTCCAAGACCAACAGAG	1474
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	1475	CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAG-----	1516
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	1517	-----ATCTCTCTGGG	1528
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle	300
Db	1529	CTCACCCCTGCGCCTGTGGAGCGTGTATCTGTTAGAAGCGAACAGAGCGTTGATGCCGATA	1588
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	1589	ACAAGATCGCCTTAAGTTACAGCAGAGAGCGCCTCACGAAGACGTCCAGGTGTGGCCG	1648
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1649	TGGGCACGTTTTCGAACCGGTTCTGTGATACCTGGGCCAAGGATGAGGACACTGTGCTC	1708
QY	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1709	AAGCATCTTAGGGCTTATGAAAGAACTAACAAAGAAAGCAGGGGAGACTGCACCCCA	1768
QY	361	AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys	380
Db	1769	GCCAAACCCGAGCAAGGCTCGTCGCATCCAGGCTGTGCCGCTTCACGTCGGCGGAAG	1828
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1829	ACCCTCTGCAAGGGGGAAGGCAAGGCCCTCCAGGCCACCAAGCCCGGTTCCGCGGCC	1888
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla	420
Db	1889	ATTGTGTCAGCTTCCCGCCACGCGCACCTCGTCTTCCACACCTGTCTGTGGGGCT	1948
QY	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1949	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGCCCCAGCCGCGCTGCTCAG	2008
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	2009	GGAGGACCTCAGGGTTCCTGAGATTCTTGCAGTGAATCTCATGCCCCGCTCCCAACG	2068
QY	461	AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
Db	2069	GACCTGAGCGTAGAGGGCCCTTGTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC	2128
QY	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu	500
Db	2129	CGTGCCATATCCCAAGAGACCAAGCAGTGGCCCTGCTGGCAGGCTGAACACCTCGGAG	2188
QY	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520

Db 2189 CCGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCCAGGGCACCCTTCAGA 2248
QY 521 AAlaArgSPGluGlnProCysAlaProThrSerGlyProCysLeuCySGlyLeuHisLeu 540
Db 2249 GCTAGGAGACGAACAGCGGTGTGCTCCCACTCAGGCGCTTGCTGCGGCGCTCCACTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335

RESULT 7

US-10-094-466-61
; Sequence 61, Application US/10094466
; Publication No. US2003020363A1
; GENERAL INFORMATION:
; APPLICANT: Spytek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 61
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)..(1678)
US-10-094-466-61

Alignment Scores:

Pred. No.: 4.67e-247 Length: 1752
Score: 2784.00 Matches: 515
Percent Similarity: 96.2% Conservative: 13
Best Local Similarity: 93.8% Mismatches: 21
Query Match: 93.3% Indels: 0
DB: 6 Gaps: 0

US-10-071-838-2 (1-549) x US-10-094-466-61 (1-1752)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 31 ATGGACGTGCTAGAGGTCGCGGCTAGTTGGTGGGCACAAGAGCGAGGACATCATATATG 90
QY 21 LysTyGlyLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysPhePheArg 40
Db 91 AATATCGAAAGGAGACACGAGCTGGGCTGCCAGAGACAAAGGGCGCTTAAGTCTTTTGA 150
QY 41 SerTyAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60

Db 151 AGCTACAAACAACAGCTGCATCATTTGGGATTTGTACAGAGACGAGCTGCTCTCTG 210
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 211 ACTGCGCGGAGGTGAAGCAAAATTCGCGGGAGATCAAGCGAAAGACAAAGTGGTGAAA 270
QY 81 MetLeuGlyAspTrpGlyLysTyLysSerSerArgLysLeuIleAspArgAlaTyLys 100
Db 271 ATGCTGGAGAAATGGACACCTTACAAAACAGACAAAGAACTCATAGATCGAGCGTACCAG 330
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 331 GGAATTCATGAAACATCCGGGGCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATC 390
QY 121 LysLeuLysAsnProGlyArgTyGlnIleMetLysGlyLysGlyLysArgSerSerGlu 140
Db 391 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCAATCTGAA 450
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 451 CACATCCAGACAGATGACCTGAGCTGAAGCGGACATTAAGAGCATATATTCTTCAGG 510
QY 161 AspArgTyGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyGlyGluTyLys 180
Db 511 GATCATACGGAACCAAGCAGCGGAACTACTTACATCTCTCGCGTATGAGAGTAT 570
QY 181 AsnProGluValGlyTyLysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyLys 200
Db 571 AACCCGAGGTGGCTACTGCAAGGAACTGAGCCACATCGCCCTTGTCTCTCTTAT 630
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 631 CTTCTGAGAGATGATCATTTGGGCACTGTGCAAGCTGTGCGCAGTGAAGAGCACTCC 690
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 691 CTGCAAGGTTTTCACAGCCCAATGGCGGAGCCGTCCAGGGCTCCAAAGACCAAGAGAG 750
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 751 CATGTGTAGCCACGTCACACTACCCAAACCATGTGGCATTCAGAGACAAAGAAATCTATGT 810
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 811 GGGCAGTGTTCCTTCAAGCTGCTCATCCGATATTGATGACGGGATCTCTCCGGG 870
QY 281 LeuThrLeuArgLeuTrpAspValTyLysValGluGlyGlnAlaLeuMetProIle 300
Db 871 CTCACCTGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAACAAGCGTTGATGCCGATA 930
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 931 ACAAGATCGCTTTAAGTTCAAGCTAGACGCGCTCACGAAGACGTCAGGTGTGCCCG 990
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 991 TGGGCAAGTTTGGAAACCGGTTCGTGATGCCGTGGCGAGGATGATGACACTGTGCTC 1050
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1051 AAGCATCTTAGGCGCTCTATGAAGAACTAACAAGAAAGGAGGAGCTGCCACCCCA 1110
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1111 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCGCTGTGCCACTTCACGTGGCGGGAAG 1170
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1171 ACCCTCTGCAAGGGGAGACAGGACGCCCTCCAGGCCACCAAGCCGGTTCCCATGGCCC 1230
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420

Db 1231 ATTGGTCAGCTTCCCCGCCACGGGCACCTCGTCTTCCACACCTGTCTGTGGGCT 1290
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1291 GTCCGGGAAGACACCTACCTGTGGCACTCAGGGTGTGCCCAAGCCCGGCTGTGAG 1350
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1351 GGAGGACCTCAGGGTTCCTGGAGATTCTGCAGTGAACCTCAGCCCGCTCCCAAG 1410
Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1411 GACCTGGACGTAGGGGACCTTGCTGCCGCTATGATTGACAGACAGAGCTGTGGGTC 1470
Qy 481 ArgAlaIleSerGlnGlyAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1471 CGTGCCATATCCACGAGGACCAAGCCGGCCACTGTGGCAGGCTGAACACCTGCGAG 1530
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1531 CGGTGAGATCGGCTTTCAGTGCGCCACACATGATTCGACACAGGGCACCCCTTCAGA 1590
Qy 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1591 GCTAGGAGACGACAGACAGTGTGTCTCCACCTCAGGACCTTGCTGCGGCTCCACTTG 1650
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 1651 GAAAGTTCACAGTTCCTCCAGGCTTC 1677

RESULT 8

US-10-071-838-5
; Sequence 5, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1538)
US-10-071-838-5

Alignment Scores:
Pred. No.: 5.22e-246 Length: 1862
Score: 2773.00 Matches: 515
Percent Similarity: 93.8% Conservative: 0
Best Local Similarity: 93.8% Mismatches: 0
Query Match: 93.0% Indels: 34
DB: 5 Gaps: 1

US-10-071-838-2 (1-549) x US-10-071-838-5 (1-1862)

Qy 1 MetAspValValGluValAlaGlySerTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGGACGTGTAGAGGTCGCGGCGAGTTGTGGGCACAGAAGCGAGAGACATCATTTATG 60

Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGACAGAGGGCCCTAAGCCTTTTCA 120
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 121 AGCTACACACACACGTCGATCATTTGGGGATTGTACATGACACGAGCTGCTCTCTG 180
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 181 ACTCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGACAAAGTGGTGAT 240
Qy 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 241 ATGCTGGAGACTGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAG 300
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAATG 360
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAAACCCCGAAGATACCATGATGAAAGGAAAGGCGCAAGGTCTATCTGAG 420
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGGAACATATATCTTCACG 480
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
Db 481 GATCGATACGAAACCAAGCAGCGGAACTACTCCACATCTCTCGCATATGAGAGTAT 540
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIlePheLeuLeuTyr 200
Db 541 AACCCGGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTTCCTGAGGAGATGCATTCCTGGGCACTGTGTCAGTCTGTGCCAGTGAGGCACTCC 660
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 661 CTGACGAGATTTCACAGCCCAATGCGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 720
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGCACTAG----- 762
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 762 ----- 762
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 763 -----TATCTGTAGAAAGCGCAACAGCGCTTGATGCCGATA 798
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 799 ACAAGAAATCGCTTAAAGTTCAAGAGCGCCTCAAGAAAGCTCCAGGTGTGCCCG 858
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 859 TGGGCACTTTTTCACCGGTTGTTGATATACCTGGGCCAGGATGAGGACACTGTGCTC 918
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 919 AAGCATCTTAGGGCTCTATGAAGAACTAACAAGAAAGCAGGGGAGCCTGCCACCCCA 978
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 979 GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG 1038
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400

Db 1039 ACCCTCTGCAAGGGGAGACGAGGCCCTCCAGGCCACCGCCGGTTCCCGCGGCC 1098
Qy 401 ILETpSerAlaSerProProArgAlaProArgSerSerThrProCySPROGLyGlyAla 420
Db 1099 ATTGGTCAGCTTCCCCCGCCACGGGCGACTCGTTCTTCACACCTGTCTGTGGGCT 1158
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1159 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCCGCCCGGCTGGCTCAG 1218
Qy 441 GLyGLyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1219 GGAGACCTCAGGGTTCCTGGAGATTCTCGACGTGAACATGCCCGCCCTCCCAAG 1278
Qy 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCySPTrpVal 480
Db 1279 GACCTGGACGTACAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGCTGGCTC 1338
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCySPTrpGlnAlaGluHisProAlaGlu 500
Db 1339 CGTGCCATATCCAGAGAGACCACTGGCCCCCTGCTGGCAGGCTGAACACCCCTGCGAG 1398
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1399 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAGGGCACCCCTTGAGA 1458
Qy 521 AlaArgAspGluGlnProCySPAlaProThrSerGlyProCySPLeuCySPGlyLeuHisLeu 540
Db 1459 GCTAGGAGACAGACCGGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTG 1518
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 1519 GAAAGTTCACGTTCCCTCCAGGCTTC 1545

RESULT 9
US-10-450-763-21335
; Sequence 21335, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 21335
; LENGTH: 7856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1697)..(4054)
; OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number
; OTHER INFORMATION: X63546,Smith-Waterman Score=4233.
US-10-450-763-21335

Alignment Scores:
Pred. No.: 1.62e-191 Length: 7856
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 9 Gaps: 1

US-10-071-838-2 (1-549) x US-10-450-763-21335 (1-7856)

Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGGACATGCTAGAGAATGCAGATAGTTTGCAGGCAACAGAGCGGAAGACATACTTATG 1756
Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTGAGCCCGTT--- 1813
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db 1814 GGAAATCAACAGCAGCATTTGATGCTTTTGGCATTGTCATGAGACGAGCTGCTCTGTG 1873
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGACGGGAGCGGAAAGAAATTCGGCGGAGATGACACGAACGAAAGTGGATGAA 1933
Qy 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 1934 ATGCTGGAGAGATGGAGACATATAAGCACAGTAGCAAACTCATAGATCGAGTGTACAAG 1993
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GGAATTCCATGAACATCCGGGGCCCGGTGTGTGTCAGTCTCTGAAACATTCAGGAATC 2053
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAGAGGGCAAGAGTCACTGAA 2113
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACACATCGACCTGAGCGTAGAGCAGACTCTCCGAAACCATGTCTTTAGG 2173
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 2174 GATCGATATGAGCCCAAGCAGAGGAATATTCTACATCTCTCGCCTATTCCGAGTAT 2233
Qy 181 AsnProGluValGlyTyrCySPArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 2234 AACC CGAGTGGGCTACTGCAGGAGCTGAGCCACATCACCGCCTGTCTCTTAT 2293
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCTGAGAGGAGCAGCATTTCTGGGCACTGTGTGCACTGTGCGCAAGTGAAGGCACTCC 2353
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCCAGATTCCACAGCCCAATGTTGGACAGTCCAGGGGCTCCAAAGACCAACAGAG 2413
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCyS 260
Db 2414 CATGTGTACCAAGTCAACACCCAAGACCATGTGGCATTCAGGACAAAGAGGTCTATGC 2473
Qy 261 GlyGlnCySPerProLeuGlyCySPLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTGCTCGTTAGGCTGCTCTCCGAACTGATTGACGGGATCTCTCGGG 2533
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 2534 CTCACCTTCGCGCTGTGGGACGTGATTGTGGTGAAGAGAAACAGTGTGATGCCAATA 2593
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCySPGlyPro 320
Db 2594 ACCAGCATGCTCTTAAGGTTACAGCAAGCGCCTCATGAAGAATCCAGGTGTGGCTG 2653
Qy 321 TrpAlaArgPheCySPAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGCACGTGTGCGGAACCAATTCTTGATACCTGGCCATGAACGATGACACCGTGTCTC 2713
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATCTTAGGGCTCTACGAAGAACTAACAAAGGAAGCAAGGGACCTGCCACCCCA 2773

QY 401 ILeTrpSerAlaSerProProAlaAlaProAlgSerSerThrProCysProGlyGlyAla 420
Db 2894 ATTGCTCAGCTTCCCGCCATGGGCACTCTGTTTTTCCAGCCCTGTCTCTGGTGGGCT 2953
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 2954 GTCCGGGAAGACACGTACCCTGTGGCACTCAGGGTGTGCCAGCCTGGCCCTGGCTCAG 3013
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 3014 GGAGGACCTCAGGGTCTCTGAGATTCCTGAGTGAAGTCAATGCCCGGCTCCCAAG 3073
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 3074 GACCTGGATATAGGGGGCCCTTGTTCCCATTTATGATTTTGAAAGGAGCTGCTGGTTC 3133
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 3134 CGTGCATATCCAGAGAGACCAGCTGGCCACCTGTGGCAGGCTGAACACTGCCGAGAG 3193

RESULT 11
US-09-962-436-562
; Sequence 562, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Sopet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 562
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-562

Alignment Scores:
Pred. No.: 1.62e-191 Length: 7878
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 3 Gaps: 1

US-10-071-838-2 (1-549) x US-09-962-436-562 (1-7878)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGGACATGCTAGAGATGCAGATAGTTTGACGACACAGAGCGGAAGACATACCTATG 1756
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCTGAGCCCGTT--- 1813
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 1814 GGAATCAACAGCAGCAATGATCGTTTGGCAITTTGTCATGAGACGAGCTGCCTCCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGCAGCGAGCGAAGAAATTCGGCGGAGATGACACGAAACGAGCAAGTGATGAA 1933
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 1934 ATGCTGGAGAAATGGAGACATATAGCACAGTAGCAAACTCATAGATCGAGTGTAACAAG 1993

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GGAATTCCTCATGACATCCGGGGCCCGGTGTGTGTCAGTCTCTGAAACATTCAGGAATC 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAAAAACCCCGGAAGATACCATGATCATGAAGAGAGGGGCAAGAGTCACTGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCATGACCTGACCTGAGCGTGAAGACGACTCTCCGGAACCATGTCTTTTAGG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 2174 GATCGATATGAGCCCAAGCAGAGGAACTATTCTACATCTCTCGCTTATTCGAGTAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
Db 2234 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCACGCCCTGTCTCTCTTAT 2293
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCCTGAGAGAGACGCAATCTGGGCACTGTGTCAGTCTGCGCAAGTGAAGAGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCCAGATTCCACAGCCCAATGTGTGGACAGTCCAGGGGCTCCAAAGCCAAACAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 2414 CATGTGTACCCAAAGTCACAAACCAAGACCATGTGGCATCAGACAAAGAAAGTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTGCTCGTTAGGCTGCTCTCCGAACTGATTGACGGATCTCTCGGG 2533
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 2534 CTCACCTCGCGCTGTGGACGTGATTTGTGGAAGAGAAACAGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGCATTCCTCTTAAGTTACAGAAAGCGCCTCATGAAGACATCCAGGTGTGSCCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGCAGCTGTGCGGAACCAATTTTCGATACCTGGGCCATGAACGATGACACCGTCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATCTTAGGGCTCTACGAAGAACTAACAAAGGAAGCAAGGAGCTGCACCCCA 2773
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 2774 GCCAAACGCGAAGAGGTCTTGACCAAGCCCTGTGCGGCTTCACGTGTGGGAAG 2833
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 2834 ACCCTCTGCAAGGGGTATAGGCAAGGCCCTCCAGGCCCAAGCCCAAGTTCAGCGGCC 2893
QY 401 ILeTrpSerAlaSerProProAlaAlaProAlgSerSerThrProCysProGlyGlyAla 420
Db 2894 ATTGCTCAGCTTCCCGCCATGGGCACTCTGTTTTCACAGCCCTGTCTCTGTGGGCT 2953
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 2954 GTCCGGGAAGACACGTACCCTGTGGCACTCAGGGTGTGCCAGCCTGGCCCTGGCTCAG 3013
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 3014 GGAGGACCTCAGGGTCTCTGAGATTCCTGGAAGTGAAGTCAATGCCCGGCTCCCAAG 3073

QY 461 AspleuaspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
|||:::|||||
DB 3074 GACCTGGATATAGGGGCCCTTGTTCCCATATGATTTGAACGAGAGCTGGGTC 3133
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
|||:::|||||
DB 3134 CGTGCCATATCCAGAGACGAGCTGGCCACCTGCTGGCAGGCTGAACACTGCGAGAG 3193

RESULT 12

US-10-843-641A-3021
; Sequence 3021, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843, 641A
; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873, 367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954, 531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954, 456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962, 436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962, 832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964, 824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967, 768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968, 007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969, 347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969, 708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3021
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-3021

Alignment Scores:

Pred. No.: 1.62e-191 Length: 7878
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 9 Gaps: 1

US-10-071-838-2 (1-549) x US-10-843-641A-3021 (1-7878)

QY 1 MetaspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
|||:::|||||
DB 1697 ATGGACATGTGTAGAGATGCAGATAGTTTTCAGAGGCACAGAGCGGAGGACATATCTATG 1756
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
|||:::|||||
DB 1757 AAGTATGACAAAGGACACCGAGCTGGCTGCAGAGGACAAAGGGCCTGAGCCGTT-- 1813
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
|||:::|||||
DB 1814 GGAATCAACAGCAGCATGTGTTTGGCATTTTGCATGAGACGAGCTGCTCCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
|||:::|||||
DB 1874 ACTGCACGGGAGCGGAGAGAAATTCCGGCGGAGATGACAGAAAGCAAGTGGATGAA 1933

QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
|||:::|||||
DB 1934 ATGCTGGAGAGATGGGAGACATATAAGCACAGTAGCAAACTCATAGATCGAGTGTACAA 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
|||:::|||||
DB 1994 GGAATTCATGACATCCGGGGCCGGTGTGTGACTCTCTGAAACATTCAGAAATC 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
|||:::|||||
DB 2054 AAGTTGAAAAACCCGGAAGATACAGATCATGAGAGAGGGCAAGAGGTACTTGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
|||:::|||||
DB 2114 CACATCCACCATCATCGACCTGAGCCTGAGACGACACTCTCCGGAACCATGTCTTCTTAG 2173
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
|||:::|||||
DB 2174 GATCGATATGAGCCAAAGCAGAGGAACATATTCTACATCTCTGCGCTTATTCGAGTAT 2233
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
|||:::|||||
DB 2234 AACCCGAGGTGGCTACTGACGAGCACTGAGCCACATACCGCTGTGTTCTTCTTAT 2293
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
|||:::|||||
DB 2294 CTGCTGAGGAGAGAGCATCTGGGCACTGTGACGTGCTGGCCAGTGAAGGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
|||:::|||||
DB 2354 CTGCCAGATTCACACAGCCCAATGTGGAGCAGTCCAGGGGCTCCAAGACCAAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
|||:::|||||
DB 2414 CATGTGTAACCAAGTACAAACCAAGACCATGTGGCATCAGGACAAAGAGGTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
|||:::|||||
DB 2474 GGGCAGTGTGCTCGTTAGGCTGCTTCCGGAACCTGATTGACGGGATCTCTCGGG 2533
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
|||:::|||||
DB 2534 CTCACCTGCGCCTGTGGAGCGTATTTGTGGAAGAGAAACAGGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
|||:::|||||
DB 2594 ACCAGCATTTGCTTTAAGTTTCAGACAGAGCGCTTCATGAAGACATCCAGGTGTGCGCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
|||:::|||||
DB 2654 TGGGCACTGTGCGGAACTTCTTCGATACCTGGGCCCATGAACGATGACACCTGTGCTC 2713
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
|||:::|||||
DB 2714 AAGCATCTTAGGCGCTTACGAAGAACTAACAAGAAAGCAAGGGGAGCCTGCCACCCCA 2773
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLysLys 380
|||:::|||||
DB 2774 GCCAAACGCGAGCAAGGTCCTTGGCACCCAGGCTGTGCCGCTTACGTTGTTGGGAAG 2833
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
|||:::|||||
DB 2834 ACCCTCTGCAAGGGGTATAGGACAGGCCCTCCAGGCCCAACAGCCAGTTCAGCGGGCC 2893
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
|||:::|||||
DB 2894 ATTGCTCAGCTTCCCGCATGGGCACTCTGTTTTCACGCGCTGTCTGTGGGCT 2953
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
|||:::|||||
DB 2954 GTCCGGAAAGACACGTAACCTGTGGGCACTCAGGCTGCGCAGCTGCGCTGAG 3013


```
OY 441 GLYGLYProGInGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
|||
Db 3014 GGAGGACCTCAGGGTTCTCTGAGATTCCTGGAGTGAAGTCAATGCCCGGCTCCCAACG 3073
OY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
|||
Db 3074 GACCTGATATAGGGGGCCCTTGTTCCCATATGATTTTGAACGAGCTGCTGGGTC 3133
OY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
|||
Db 3134 CGTGCATATCCCAAGAGACCACTGGCCACTGTCAGGCTGAACACTGCCGAGAG 3193

RESULT 13
US-10-956-157-1259
; Sequence 1259, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: patentin version 3.2
; SEQ ID NO 1259
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1259
```

```
Alignment Scores:
Pred. No.: 1.62e-191 Length: 7878
Score: 2191.50 Matches: 408
Percent Similarity: 88.2% Conservative: 33
Best Local Similarity: 81.6% Mismatches: 58
Query Match: 73.5% Indels: 1
DB: 9 Gaps: 1
```

US-10-071-838-2 (1-549) x US-10-956-157-1259 (1-7878)

```
OY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
|||
Db 1697 ATGGACATGCTAGAGATGCAGATAGTTTGCAGGCACAGAGCGGAAGACATACCTATG 1756
OY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
|||
Db 1757 AAGTATGACAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTGAGCCCGTT-- 1813
OY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
|||
Db 1814 GGAATCAACAGCAGCATGATCGTTTGGCATTTTGCATGAGACGAGCTGCCCTCTGTG 1873
OY 61 ThrAlaArgGluAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80
|||
Db 1874 ACTGACGCGGAGCGAAGAAATTCGGCGGAGATGACACGAAAGCAAGCAAGTGATGAA 1933
OY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
|||
Db 1934 ATGCTGGAGAAATGGAGACATATAGCAGACAGTAACTCATAGTCAGTACAAG 1993
OY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
|||
Db 1994 GGAATTCCTCATGAACATCCGGGGCCCGGTGTGTCAGTCTCTCTGAACATTCAAGAAATC 2053
OY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
|||
Db 2054 AAGTTGAAAAACCCCGAAGATACCATGATCAATGAAGAGAGGGGCAAGAGTCACTGAA 2113
OY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
|||
Db 2114 CACATCCACCAACATGCACTTGAGCGTGAAGAGACTCTCCGAAACCATGTCTTCTTAGG 2173
```

```
OY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
|||
Db 2174 GATCGATATGGAAGCCAAAGCAGAGGGAACATATTCTACATCTCCTCGCCTATTCGAGTAT 2233
OY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
|||
Db 2234 AACCCGGAGGTGGCTACTGCAAGGACCTGAGCCACATCACCGCCTGTCTCTTAT 2293
OY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
|||
Db 2294 CTGCTGAGGAGGACGCAATTCTGGGCACTGGTGCACTGCTGGCCAAGTGAAGGCACTCC 2353
OY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
|||
Db 2354 CTGCAAGGATTCACAGCCCAATGTGGGACAGTCCAGGGGCTCCAAGACCAACAGGAG 2413
OY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
|||
Db 2414 CATGTGTACCCAAAGTACAAACCAAGACCATGTGGCATCAGGACAAAGAGTCTATGC 2473
OY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
|||
Db 2474 GGGCAGTGTGCTCGTTAGGCTGCTCTCCGGAACCTGATGACGGGATCTCTCGGG 2533
OY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle 300
|||
Db 2534 CTCACCTTGCGCTGTGGAGCGTGTATTGTGTGAAGAGAAACAGGTGTGATGCCAATA 2593
OY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
|||
Db 2594 ACCAGCATGTCTTTAAGGTTACAGAGAGCGCTCATGAAGACATCCAGGTGTGGCTTG 2653
OY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
|||
Db 2654 TGGGACGTCCTCGGAACCAATTCTTGATACCTGGGCGATGAACGATGACACCGTCTC 2713
OY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
|||
Db 2714 AAGCATCTTAGGGCTCTAAGGAACTAACAAGAGCAAGGAGGAGCTGCCACCCCA 2773
OY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
|||
Db 2774 GCCAAACGGAGCAAGGGTCTTGGCAACCCAGGCGCTGTGCCGCTTCACTGTGGGAAG 2833
OY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
|||
Db 2834 ACCCTCTGCAAGGGGTATAGGAGGCCCTCCAGGCCCCACCAAGCCAGTTCCAGCGGCC 2893
OY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyAla 420
|||
Db 2894 ATTTGCTGACTTCCCGGCCATGGGCATCTCGTTTTCACAGCCCTGTCTGTGGGGCT 2953
OY 421 ValArgLysAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
|||
Db 2954 GTCCGGAGACACAGTACCCCTGTGGCACTCAGGGTGTGCCACCTGAGCCCTGGCTCAG 3013
OY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
|||
Db 3014 GAGGACCTCAAGGTTCTCTGAGATTCTTGAGTGAAGTCAATGCCCGGCTCCCAACG 3073
OY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
|||
Db 3074 GACCTGATATAGGGGGCCCTTGTTCCCATATGATTTTGAACGAGGCTGCTGGGTC 3133
OY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
|||
Db 3134 CGTGCATATCCCAAGAGACCACTGGCCACTGTCAGGCTGAACACTGCCGAGAG 3193

RESULT 14
US-10-450-763-24016
; Sequence 24016, Application US/10450763
; Publication No. US20050196754A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 24016
; LENGTH: 8180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (4369)..(6126)
; OTHER INFORMATION: 97% homologous to Homo sapiens oncogene, accession number
; US-10-450-763-24016

Alignment Scores:
Pred. No.:      3.88e-186      Length:      8180
Score:          2133.50       Matches:      408
Percent Similarity: 77.2%      Conservative: 33
Best Local Similarity: 71.5%      Mismatches:  58
Query Match:    71.5%          Indels:       72
DB:             9              Gaps:         2

US-10-071-838-2 (1-549) x US-10-450-763-24016 (1-8180)

QY      1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB      1697 ATGCACATGCTAGAGAATGCAGATAGTTTGCAGGCACAGACGGAGACATACCTATG 1756

QY      21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB      1757 AAGTATGACAAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCCTGAGCCCGTT-- 1813

QY      41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB      1814 GGAATCAACAGCAGCATTTGATCGTTTGGCATTTTGCATGAGACGGAGCTGCTCTGTG 1873

QY      61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB      1874 ACTGCACGGGAGCGGAGAAATTCCGGCGGAGATGACACGAACGACGAAGTGTGAA 1933

QY      81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB      1934 ATGCTGGGAGAAATGGAGACATATAAGCACAGTAGCAAACTCATAGATCGAGTGTACAA 1993

QY      101 GlyMetCProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB      1994 GGAATTCCTCATGAACATCCGGGGCCCGGTGTGTCAGTCTCTGAAACATTCAGAAATC 2053

QY      121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB      2054 AAGTTGAAAAAACCCTCGAAGATACCAAGATCATGAAGAGAGGGGCAAGAGTCATCTGAA 2113

QY      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB      2114 CACATCCACACACATGCACCTGAGCGTGAAGACGACGCTCTCCGGAACCATGCTTCTTAAG 2173

QY      161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB      2174 GATCGATATGAGCCAAAGCAGAGGAACATATTCTACATCTCTGCTGCTATTCCGAGTAT 2233

QY      181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
```

```
DB      2234 AACCCGAGGTGGGCTACTGCAGGAGCCTGAGCCACATCACCGCCTTGTCTCTTAT 2293
QY      201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB      2294 CTGCCTGAGGAGAGCAGCATTTCTGGGCACTGTGCAGCTGCTGGCCAGTGAGAGGCACTCC 2353
QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB      2354 CTGCCAGGATTCACAGCCCAATGTGGGACAGTCCAGGGGCTCCAAAGACCAACAGAG 2413
QY      241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB      2414 CATGTGTACCCAAAGTCACAAACCAAGACCATGTGGCATCAGACCAAGAAAGTCTATGC 2473
QY      261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB      2474 GGGCAGTGTGCTCTGTTAGGCTGCTTCTCCGGAACCTGATGACGGATCTCTCGGG 2533
QY      281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB      2534 CTCACCCCTGGCCCTGTGGAGCGTATTTGTGGAAGAGAAACAGGTGTGATGCCAATA 2593
QY      301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB      2594 ACCAGCATTTGCTCTTAAGGTTCAGACAGAAAGCGCTCATGAAAGCATCCAGGTGTGGCTG 2653
QY      321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB      2654 TGGGACGCTCTCGGAACCAATTCTTCGATACCTGGGCCATGAACGATGACACCGTGTCTC 2713
QY      341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB      2714 AAGCATTTAGGGCTCTACGAGAAACTTAACAAGAAAGCAAGGGGACCTGCCACCCCA 2773
QY      360 ----- 360
DB      2774 GGCCCAACAGCCCTGGAGCAAGGTGTGTGCAGAAAGCCCCAGCAGTGTGAACCTTG 2833
QY      360 ----- 360
DB      2834 GGGGAGTCCCAAGAGCCACCAATGCCCCAACGGCTTCCCATGCCAGGACACACAC 2893
QY      360 ----- 360
DB      2894 ACCCTTCCTCTGGATCAGCAGACTACAGGCGTGTCTCAGTGTCAAGACCAAGGGGCC 2953
QY      361 -----AlaLysProGluGlnGlySerSerAlaSer 370
DB      2954 ACACAGAGACCCCAAGACTCCAGAGATGCAGCCAAACGAGCAAGGGTCTTGGCACCC 3013
QY      370 rArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
DB      3014 CAGGCGTGTGCCGGCTTCACTGTGTGGAGAAGCCCTCTGCAAGGGGTATAGGCAAGGCCCC 3073
QY      390 OProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPr 410
DB      3074 TCCAGGCCCCACCAAGCCCAAGTTCAGCGGCCCATTTGCTCAGCTTCCCGCCATGGGCCATC 3133
QY      410 OArgSerSerThrProCysProGlyGlyAlaValAlaArgGluAspThrTyrProValGlyTh 430
DB      3134 TCGTTTTCACAGCGCCCTGTGTGTGGGTGTCTCCGGGAAGACACGTAACCTGTGTGGCAC 3193
QY      430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLe 450
DB      3194 TCAGGGTGTGCCAGCGCTGCGCTGTGAGGAGGACCTCAGGGTTCCTGAGATTCCT 3253
QY      450 uGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheAr 470
DB      3254 GGAAGTGAAGTCAATGCCCCGGCTCCCAACGAGCTGATATAGGGGGCTTGGTCCC 3313
QY      470 gHisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGlnAspGlnLeuAl 490
DB      3314 CCAATTATGATTTTGAACGAGCTGTGGGTCTCGTGCATATCCACGAGAGACCAAGCTGGC 3373
```

Qy 490 aProCytrpGlnAlaGluHisProAlaGlu 500
Db 3374 CACCTGCTGCAGGCTGAACACTGCCGAGAG 3404

RESULT 15
US-10-283-975A-272

; Sequence 272, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283, 975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340, 938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338, 997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340, 081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341, 012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 8201
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-272

Alignment Scores:
Pred. No.: 3.9e-186 Length: 8201
Score: 2133.50 Matches: 408
Percent Similarity: 77.2% Conservative: 33
Best Local Similarity: 71.5% Mismatches: 58
Query Match: 71.5% Indels: 72
DB: 7 Gaps: 2

US-10-071-838-2 (1-549) x US-10-283-975A-272 (1-8201)

Qy 1 MetAspValValGluValAlaGlySerTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGCACATGGTAGAAGATGCAGATAGTTGCAGGACAGAGCGGAAGACATACCTATG 1756
Qy 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGACACCGAGCTGGCTGCCAAGAGCAAGGGCCTGAGCCCGTT--- 1813
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 1814 GGAATCAACAGCAGCATTTGATCGTTTGGCATTTTGCATGAGACGAGCTGCCCTCTGTG 1873
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGCACGGGAGCGGAAGAAATTCCGGCGGAGATGACACGAAAGCAAGTGGATGAA 1933
Qy 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspAlaTyrLys 100
Db 1934 ATGCTGGAGAAATGGAGACATATTAAGCACAGTAACTCATATGATCGAGTGTACAG 1993
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GGAATTCACATGACATCCGGGGCCCCGGTGGTCAATCCTCTGAACTTCAGGAATC 2053
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 2054 AAGTTGAAAAACCCCGAAGATACCATGATCATGAAGAGGCGCAAGAGTCACTGAA 2113
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCATGATGACCTGAGCGTAGAGCACTCTCCGGAACCATGTCTTCTTTAGG 2173

Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 2174 GATCGATATGAGCCCAAGCAGAGGAATATTCTACATCTCTCGGCTTATTCGAGTAT 2233
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
Db 2234 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCACCGCCTGTCTCTCTTAT 2293
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCCTGAGGAGAGCAGCATTCGGGCACTGGTGACGCTGGCCAGTGAGGCACTCC 2353
Qy 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCCAGGATTCACACAGCCCAATGCTGGGACAGTCCAGGGGCTCCAAAGACCAAGAG 2413
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 2414 CATGTGTACCCCAAGTCAACACCAAGACCATGTGGCATCAGGACAAAGAGTCTATGC 2473
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTGCTCTGTTAGGCTGCTTCCGGAACCTGATTCAGCGGATCTCTCGGG 2533
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 2534 CTCACCCCTGGCCTGTGGAGCGTGTATTGGTGGAAGAGAACAGTGTGATGCCAATA 2593
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGCATTTGCTTAAAGGTTCAGACAGAGCGCCTCATGAAGACATCCAGGTGTGGCCTG 2653
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGGCACGCTGCGGAACCAATTCTTCGATACCTGGGCCATGAACGATGACACCGTGCTC 2713
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATTTAAGGCTCTACGAAAGAACTTAACAAGGAAGCAAGGGGACCTGCCACCCCA 2773
Qy 360 ----- 360
Db 2774 GGCCCAACAGCCCTGGAGCAAGGTGTGTGCAGAAAGCCCCAGCAGTCTGAACCTGTG 2833
Qy 360 ----- 360
Db 2834 GGGGCAGTCCAGAGCAACCAATGCCCCCAAGCGCTTCCCCTATGCCAGGCAACAC 2893
Qy 360 ----- 360
Db 2894 ACCCTCCCTCTGGGATCAGCAGACTACAGGCGTGTCTCAGTGTCAAGACCAAGGGCC 2953
Qy 361 -----AlaLysProGluGlnGlySerSerAlaSer 370
Db 2954 ACACAGAGACCCCAAGGACTCCAGAGATGCAGCCCAACCGGAGCAAGGTCCTTGGCACC 3013
Qy 370 rArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
Db 3014 CAGGCTGTGCGGCTTCACTGTGTGGAAAGACCTCTGCAGAGGGGTATAGGCAGGCCCC 3073
Qy 390 oProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPr 410
Db 3074 TCCAGGCCCAACAGCCAGTTCAGCGGCCCATTTGCTCAGCTTCCCGCCATGGGCATC 3133
Qy 410 oArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyTh 430
Db 3134 TCGTTTTCACAGCCCTGTCTGTGTGGGGCTGTCCGGGAAGACAGTACCTGTGGGCAC 3193
Qy 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTyrArgPheLe 450
Db 3194 TCAGGGGTGTGCCAGCCTGGCGCTGGCTCAGGGAGGACCTCAAGGTTCTCTGAGATTCTCT 3253
Qy 450 uGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheAr 470

Db 3254 GGAGTGAAGTCAATGCCCGGCTCCCAACGACCTGATATAGGGGCCCTTGTTCCC 3313
QY 470 GHISTYRASPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAl 490
Db 3314 CCATTATGATTGTAACGAGCTGTGGTCCGTGCCATATCCAGAGACCAAGCTGGC 3373
QY 490 aProCysTrpGlnAlaGluHisProAlaGlu 500
Db 3374 CACCTGCTGGCAGGCTGAACACTGCGGAGAG 3404

Search completed: April 5, 2006, 15:15:55
Job time : 1256 secs

APR 10 2006 16:21:15

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:49:32 ; Search time 289 Seconds
(without alignments)
3376.752 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVVEVAGSMWAQERDILM.....TSGPCLCGHLHSSQFPFGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/abs8/ABSSWEB_spool/US10071838/runat_05042006_141917_5227/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=tni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs807 -USER=US10071838 @CGN_1_1_193 @runat_05042006_141917_5227 -NCPU=6
-ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/1_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2850	95.5	2084	3	US-09-799-451-692 Sequence 692, App
2	2827	94.8	2072	3	US-09-799-451-691 Sequence 691, App
3	2133.5	71.5	8201	2	US-08-253-155A-9 Sequence 9, Appl1
4	344	11.5	1981	3	US-09-620-312D-715 Sequence 715, App
5	337	11.3	1823	3	US-09-774-528-336 Sequence 336, App
6	337	11.3	1823	3	US-10-120-988-336 Sequence 336, App
7	332.5	11.1	1925	3	US-09-949-016-2139 Sequence 2139, Ap
8	258	8.6	3039	3	US-09-620-312D-675 Sequence 675, App
9	253	8.5	2955	3	US-09-620-312D-676 Sequence 676, App

10	238	8.0	1083	3	US-09-248-796A-5457	Sequence 5457, Ap
11	226	7.6	2456	3	US-09-949-016-1620	Sequence 1620, Ap
12	222	7.4	1887	3	US-09-270-767-12482	Sequence 12482, A
13	194.5	6.5	806	3	US-09-270-767-9899	Sequence 9899, Ap
14	182.5	6.1	4039	2	US-08-363-300-1	Sequence 1, Appli
15	175.5	5.9	3983	3	US-09-762-311-3	Sequence 3, Appli
16	175.5	5.9	3988	3	US-09-762-311-4	Sequence 4, Appli
17	172	5.8	716	3	US-09-533-559-4537	Sequence 4537, Ap
18	172	5.8	1056	3	US-09-248-796A-1133	Sequence 1133, Ap
19	170.5	5.7	2574	3	US-10-104-047-400	Sequence 400, App
20	169	5.7	2066	3	US-10-104-047-56	Sequence 56, Appl
21	164	5.5	20113	3	US-09-902-540-1173	Sequence 1173, Ap
22	163.5	5.5	58857	3	US-09-477-962-1	Sequence 1, Appl1
23	162	5.4	61158	3	US-09-949-016-15041	Sequence 15041, A
24	162	5.4	75431	3	US-09-949-016-15122	Sequence 15122, A
25	158.5	5.3	9818	3	US-09-902-540-987	Sequence 987, App
26	156	5.2	17315	3	US-09-902-540-1103	Sequence 1103, Ap
27	155	5.2	30783	3	US-09-902-540-1258	Sequence 1258, Ap
28	154	5.2	1659	3	US-09-248-796A-1820	Sequence 1820, Ap
29	151.5	5.1	5418	3	US-09-949-016-2477	Sequence 2477, Ap
30	150.5	5.0	1902	3	US-09-902-540-8758	Sequence 8758, Ap
31	150.5	5.0	9556	3	US-09-902-540-929	Sequence 929, App
32	149	5.0	3073	3	US-09-902-540-643	Sequence 643, App
33	149	5.0	7527	3	US-09-252-991A-71	Sequence 71, Appl
34	148.5	5.0	1708	3	US-10-104-047-1901	Sequence 1901, Ap
35	148.5	5.0	2502	3	US-10-190-967-8	Sequence 8, Appli
36	148.5	5.0	2631	3	US-09-902-540-4168	Sequence 4168, Ap
37	148.5	5.0	3763	3	US-09-919-039-243	Sequence 243, App
38	148.5	5.0	5185	3	US-09-976-594-640	Sequence 640, App
39	148.5	5.0	18686	3	US-09-902-540-1206	Sequence 1206, Ap
40	148.5	5.0	20757	3	US-09-902-540-1189	Sequence 1189, Ap
41	148.5	5.0	30780	3	US-09-902-540-1243	Sequence 1243, Ap
42	148	5.0	4800	3	US-09-902-540-562	Sequence 562, App
43	148	5.0	26012	3	US-09-902-540-1212	Sequence 1212, Ap
44	148	5.0	29272	3	US-09-902-540-1217	Sequence 1217, Ap
45	147.5	4.9	723	3	US-09-270-767-1491	Sequence 1491, Ap

ALIGNMENTS

RESULT 1
US-09-799-451-692
Sequence 692, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungqing
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Yang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163) .. (1755)
US-09-799-451-692

Alignment Scores:

Pred. No.:	5.03e-183	Length:	2084
Score:	2850.00	Matches:	528
Percent Similarity:	96.4%	Conservative:	1
Best Local Similarity:	96.2%	Mismatches:	2
Query Match:	95.5%	Indels:	18
DB:	3	Gaps:	1

US-10-071-838-2 (1-549) x US-09-799-451-692 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGTGTAGAGGTCCGGGCGAGTTGGTGGCACAAGACGAGGACATCATTAATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAGAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCTTAAGCCTTTTGA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGlyThrGluLeuProProLeu 60
Db 283 AGCTACAACAACACCTGCATCTTGGGATGTACATGAGACGGAGCTGCCTCTGTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGTCAGCGTACAAAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGGACGTAAGCGGACATTAAGGAAGCATATATTCTTCAAG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 589 GATCGATACGGAACCAAGCAGCGGGAAGTACTCCACATCTCTCGGCATATGAGGAGTAT 648
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 649 AACCCGGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTCTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCTGAGAGGATGCATTTCTGGGCACGTGTGCACTGCTGCGCAGTAGAGGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTGACAGGATTTCAAGCCAAATGCGGACGCTCCAGGGGCTCCAAGACCAAGAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCAACCAAGACCATGGGCGCATCAGACCAAGAAATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGGAATTGATTCAGCGGATCTCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle 300

Db 949 CTCACCCCTGCGCCTGTGGGACGTGTACTGTGTAGAAGCGCAACAGCGCTTGATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATCGCCTTTAAAGTTCAAGCAGCAAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 1068
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGCACGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGAGTAGGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGCTCTATGAAGAACTAACAGAAAGCAGGGGGGACCTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1189 GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGAGACAGGAGCCCTCCAGGCCACAGCCCGGTTCCCGCGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1309 ATTTGGTCAGCTTCCCCCGCACGGGCACTGTTCTTCCACACCCTGTCTGTGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACCTTACCCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1429 GGAGACCTCAGGCTTCTCGAGATTCCTGACGTGAATTCATGCCCGCTCCCAAG 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1489 GACCTGACGTAGAGGGCCCTTGGTCCGCATTATGATTTCAGACAGAGCTGTGGCTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1549 CGTGCATATATCCAGAGGAGACCAAGCTGCCCCCTGTGGCAGCTGAACACCCCTGGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CCGGTGAGATCGGCTTTCGTGCACACCACTGATTCGACCAAGGCGACCCCTTACGA 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1669 GCTAGGAGCAGACAGCGGTGTGCTCCACCTCAGGGCCTTGCTTGGCGCTCCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1729 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1755

RESULT 2
US-09-799-451-691
; Sequence 691, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei

APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 691
LENGTH: 2072
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163)..(1743)
US-09-799-451-691

Alignment Scores:
Pred. No.: 1,77e-181 Length: 2072
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 95.6% Mismatches: 2
Query Match: 94.8% Indels: 22
DB: 3 Gaps: 1

US-10-071-838-2 (1-549) x US-09-799-451-691 (1-2072)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGCTGTAGAGGTCCGGCGGACGTGTGGGGCAACAAGCGAGAGACATCATATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTCA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACAAACAACAGTCATCATTTGGGATTGTACATGAGACGAGCTGCCCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGCGAAGCAAAATTCCGCGGAGATCAGCCGAAGAGCAAGTGCGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAAGAGCTCATATGAGCGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAAATCCGGGGCCCGATGTGTCAGTCCCTCTGAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTGAATAAACCCCGGAAGATACCAATCATGAAGAGAGGCAAGGTCATCTGAG 582
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 583 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAAGAGCATATATTCTTCAG 642
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLys 180
Db 643 GATCGATACGGAACCAAGCAGCGGAATCTCCACATCCTCTGGCATATGAGAGATAT 702
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLys 200
Db 703 AACCCGAGGTGGCTACTGCAAGGAGCTGAGCCACATCGCCCTTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

Db 763 CTTCCTGAGAGAGATGCAATTCTGGGCACTGTGTGACGCTGCTGGCCAGTGAGAGCACTCC 822
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 823 CTGACGGATTTCAACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
Db 883 CATGTGTAGCCAGCTCACAAACCAAGACCATGGGGCATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCTGGG 936
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 937 CTCACCTGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAAGCGGCTTGATGCCGATA 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACAAGATCGCCTTTAAGTTCAAGCAAGAGCGCTCACGAAGACGTCCAGGTGTGCCCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1057 TGGGCACTTTTTCACACCGGTTCTGTATACCTGGGCCAGGAGTGAAGACACTGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1117 AAGCATCTTAGGGCCTTATGAAAGAACTTAACAAGAAAGCAGGGGGAACCTGCAAACCCCA 1176
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1177 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGTCCGCTTACGTCGGGGAAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTGTCAAGGGGAGCAGGAGGCGCTCCAGGCCACAGCCCGGTTCCCGCGGCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1297 ATTTGTGAGCTTCCCCCGCACGGGACCTGTTCTTCCACACCTGCTGTTGGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1357 GTCCGGGAAGACCTTACCCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1416
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1417 GGAGACCTCAGGGTTCCTGAGATTCTGCAGTGGAATCCATGCCCGCTCCCAAG 1476
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1477 GACCTGAGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTCAAGACAGAGTGTGGGTC 1536
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 1537 CGTGCCATATCCCAAGAGAGACCAAGTGGCCCCCTGCTGGAGGCTGAACACCCGCGAG 1596
QY 501 ArgValArgSerAlaPheAlaIleProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1597 CGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCAACCCCTTCAAG 1656
QY 521 AlaArgAspGluGlnProCysAlaIleProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1657 GCTAGGAGCAACAGCGGTGTGTCCCACTCAGAGGCGCTTGCTGTGCGGCTTCACTTG 1716
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1717 GAAAGTTCTCAGTTCCTCCAGGCTTC 1743
RESULT 3
US-08-253-155A-9
; Sequence 9, Application US/08253155A

```
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeno
; TITLE OF INVENTION: CDK4 Binding Proteins
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-253-155A-9
```

```
Alignment Scores:
Pred. No.: 6.2e-134 Length: 8201
Score: 2133.50 Matches: 408
Percent Similarity: 77.2% Conservative: 33
Best Local Similarity: 71.5% Mismatches: 58
Query Match: 71.5% Indels: 72
DB: 2 Gaps: 2
US-10-071-838-2 (1-549) x US-08-253-155A-9 (1-8201)
```

```
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1697 ATGCACATGCTAGAGAAATGCAAGATAGTTTGACAGCACAAGAGCGGAAGACATACCTTATG 1756
QY 21 LysTYTGluLysGlyHisValArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 1757 AAGTATGACAAAGGACACCGAGCTGGGCTGCCAGAGCAAGAGGGGCTGAGCCCGTT--- 1813
QY 41 SerTYrAsnAsnAsnValAspHisValLeuGlyIleValHisGluThrGluLeuProPheLeu 60
Db 1814 GGAATCAACAGCAGCATGATGCTTTTGGCAATTTTGCATGAGACGAGCTGCCCTCTGTG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 1874 ACTGCACGAGGAGCGAAGAAATTCGGCGGAGATGACACGAAAGCAAGCAAGTGATGAA 1933
QY 81 MetLeuGlyAspTrpGluLysTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100
Db 1934 ATGCTGGAGAAATGGAGACATATAGCAGACAGTAACTCATAGATCGAGTGTAACAG 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1994 GGAATTCCCATGAACATCCGGGGCCCGGTGTGTCAAGTCTCTCTGAACATTCAAGAAATC 2053
QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
```

```
Db 2054 AAGTGA AAAACCCCGAAGATACCAATCATGAAGAGAGGGGCAAGAGTCAATCGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 2114 CACATCCACCACCATCGACCTGGACGTGAGGACGAGACTCTCCGGAACCATGTCTTAAAG 2173
QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db 2174 GATCGATATGAGCCCAAGCAGAGGAACTATTCTACATCTCTGCGCTATTCCGAGATAT 2233
QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTYr 200
Db 2234 AACCCGAGGTGGCTACTGACAGGACCTGAGCCACATCACCGCTGTCTCTTAT 2293
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 2294 CTGCGTGAAGAGACGCAATTCGGGCACTGTGCAGCTGCTGGCCAGTGAGAGGCACTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCCAGGATTCACAGACGCCAAATGGTGGAGACAGTCCAGGGGCTCCAAGACCAAGAGAG 2413
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 2414 CATGTGTTACCAAGTACAAACCAACCATGTGGCATTCAGGACCAAGGAAGGTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTGCTCGTTAGGCTGCTTCTCCGGAACCTGATTGACGGGATCTCTCTGGG 2533
QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 2534 CTCACCCCTGGCCTGTGGACGTGTATTGTGTGAAGAGAAACAGGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 2594 ACCAGCATTCCTTTAAGGTTACAGAAAGCGCTCATGAGACATCCAGGTGTGCGCTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 2654 TGGGACAGTTCGGGAAACCAATTCATGATACCTGGGCAATGAACATGACACCGGTCTC 2713
QY 341 LysHisLeuAlaArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 2714 AAGCATTTAGGGCTCTACGAAGAACTAACAAAGGAAGCAAGGGACCTGCCACCCCA 2773
QY 360 ----- 360
Db 2774 GGGCCAAAGCCCTGGAGCAAGGTGTGTGGCAGGAAGCCCCCAGCAGTCTGAACCTTG 2833
QY 360 ----- 360
Db 2834 GGGGAGTCCAGAGGACCAACCCATGCCCAACGGCTTCCCATGCCAGGACAGCAC 2893
QY 360 ----- 360
Db 2894 ACCCCTCCCTGGGATACAGACACTACAGCGGTGTCTGATGTCAGACCAAGGGGCC 2953
QY 361 ----- 361
Db 2954 ACACAGAGACCCCAAGGACTCCAGAGATGCAAGCCAAACGCGAAGGAGGTCCTTGGCACC 3013
QY 370 rArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
Db 3014 CAGGCTGTGCGGCTTCAAGTGTGGGAAGACCTCTTGAAGGGGTATAGGCGAGGCCCC 3073
QY 390 oProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaPr 410
Db 3074 TCCAGGCCCAACCAAGCCAGTTCCAGCGGCCCATTTGCTCAGCTTCCCGCCATGGGCAATC 3133
QY 410 oArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTYrProValGlyTh 430
```


Db 3134 TCGTTTTCACGCCCTGTCCTGTGGGCTGTCCGGAGAACAAGTACCTGTGGGCAC 3193
Qy 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeu 450
Db 3194 TCAGGAGTGTGCCACCTGCGCTGGCTCAAGGAGACCTCAGGGTTCTTGAGATTCTC 3253
Qy 450 uGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTrpPheAr 470
Db 3254 GGAGTGAAGTCAATGCCCGCGCTCCCAACGACCTGATATAGGGGCCCTTGTTCC 3313
Qy 470 gHisTrpAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAl 490
Db 3314 CCATTATGATTTCGAACGAGCTGCTGGGTCCGTCCATATCCAGAGAGACCAAGCTGGC 3373
Qy 490 aProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 3374 CACCTGCTGCAGGCTGAACACTGCGGAGAG 3404

RESULT 4

US-09-620-312D-715
Sequence 715, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aldong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yunging

APPLICANT: Wang, Dunrui

APPLICANT: Wang, Zhiwei

APPLICANT: John Tillinghast

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: No. 6569662el Nucleic Acids and

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1105

SOFTWARE: pt_fl_genes Version 1.0

SEQ ID NO 715

LENGTH: 1981

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (59) ..(1585)

US-09-620-312D-715

Alignment Scores:

Pred. No.: 4.51e-14 Length: 1981

Score: 344.00 Matches: 150

Percent Similarity: 40.2% Conservative: 77

Best Local Similarity: 26.5% Mismatches: 207

Query Match: 11.5% Indels: 131

DB: 3 Gaps: 26

US-10-071-838-2 (1-549) x US-09-620-312D-715 (1-1981)

Qy 8 GlySerTrpTrp-----AlaGlnGluArgGluAspIleIleMetIleTyrGlnIuLys 24
Db 46 GGGCGCGGCTGGATGCGAAGACAAACGAGAGAATGGGCCCGCGCGCGCGCGCGG 105

Qy 25 GlnHisArgAlaGlyLeuProGlnAspLysGlyProLysProPheArg-SerTyrAsnAs 44
Db 106 GGAAGCCTGTCCGGAACCCGGAGAGCCTGGCCAGAGGCCCGCAAGCCGCAACCAACCA 165
Qy 44 nasValAspHisIleuGlyIleValHisGluThrGluLeuProLeuThrAlaArgG1 64
Db 166 CGAACTCAGCTCTCTCGGG-----TCTGACTCGAGGCCAAGCCTTCGCCGAGCCCG 219
Qy 64 u-----AlaIleGlnI1 68
Db 220 CATGCAAGTTCGGCTTCATCGTGGGCTCCGAGGCGCGGAGGCGCGCTGAGAGAAGT 279
Qy 68 eArgArgGluIle-----SerArgLysSerLysTrpValAspMetLeuGlyAspTrpG1 86
Db 280 ACCCTGGAGGTCTGAGGCAAGAGGAGTCCAAAGTGGCTGACATGCTCAACAACACTGG 339
Qy 86 uLysTyr-----LysSerSerArgLysLeuIleAspArgAlaTyrLysGlyMetProMe 104
Db 340 CAATGATGAGCCCAAGAACAAAGATTGCTGTGCGGTGCCAAAAGGCGATCCGCC 399
Qy 104 tAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGluMetLysLeuLysAs 124
Db 400 TTCTTGGCGGCGCGCTGCTTGGCAGTACCTGTCAAGAGGCAAGTGAAGTTACAGCAGA 459
Qy 124 nProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlnHisIleGlnAr 144
Db 460 CCCTGGAAGTTTGACGAGCTG---GACATGTCCCTGGGAGCCCAAGTGGCTGACGT 516
Qy 144 gIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrG1 164
Db 517 GATTGAGCTGACCTGCACCCGCGAGTCCCATTCATGAATAATTGTGTCCCGGGGGGG 576
Qy 164 yThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyrAsnProGluVa 184
Db 577 CCACGGCCAGCAGACCTATTCTGCTGTGAAGGCTTACACGCTGTACCGCGCCGAGGA 636
Qy 184 lGlyTyrCysArgAspLeuSerHisIleAlaValLeuPheLeuLeuTyrLeuProGluG1 204
Db 637 GGGCTACTGACAGGCCAGGCCCATTTGCCGCTGTGCTCATGATATGCTGCTGA 696
Qy 204 uAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPh 224
Db 697 GCAAGCCTTCTGTGCTGCTGTAAGATC-----TGTGAGAAGTAC---CTGCCCGGCTA 747
Qy 224 eHisSerProAsnGlnGlyThrValGlnGlnGlyLeuGlnAspGlnGlnGlnHisIleValAla1 244
Db 748 CTACAGCGAAGAAA-----CTGAGCGGATCCAGCTGACGGGAGATCCTTTTCTC 798
Qy 244 aThrSerGln-----ProLysThrMetGlyHisGlnAspLysLysAspLeuCysG1 261
Db 799 GCTGTTCAGAAAGGTGTGCGCGGTGGCCACAAAGCACTCAAGCGTCAAGATC----- 853
Qy 261 yGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLe 281
Db 854 -----GACCCGCTCTTATATGACAGAATGTTCAATGTGCGCTTCCGAAACCTT 906
Qy 281 u-----ThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLe 297
Db 907 GCCCTGAGCTCTGTGCTGCTGCTGACATGTTCTTGTGAAGG----- 955
Qy 297 uMetProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerAr 317
Db 956 -----GTCAAGATCATCTTCCGGGTGGGGCTG----- 982
Qy 317 gCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAs 337
Db 982 ----- 982
Qy 337 pThrValLeuLysHisIleuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLe 357
Db 983 -GTGCTGCTGAAGCAGCGCGCTGGGCTCCCTGAGAAAGGTCAAAAGCTTGCACAGGCGCAGTA 1041

QY 357 uPro-ProAlaAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerA 377
Db 1042 CGAGACCATCGAGC-----GACTGCGGAGCCTCAGCCCAAGATCA----- 1082
QY 377 rgGlyGlyLeuThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgP 397
Db 1083 -----TGCAAG----- 1088
QY 397 heProArgProIleTrpSer-----AlaSerPro-ProArgAlaProArg--- 411
Db 1089 -----AGGCCTTTCGTGTCCAGAGAGTGTGTGAAGTGTCCCGTGCACAGAGCGCAGATGA 1143
QY 412 SerSerThrProCysProGlyGlyAlaValArgGluAspThrTyRProValGlyThrGln 431
Db 1144 GCGCGAACACTTCATTCAGCTGCGGCGCTGGCAGAGACCCG-----GGGTAGCT 1194
QY 432 GlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGln 451
Db 1195 GCAGTGCCTGCTCCCGCCAGGCT--GCATGTGCCAAGGCTTACTTGATGCAGAAC 1251
QY 452 TrpAsnSerMetProArgLeuProThrArgPheLeuAspValGluGlyProTrpPheArgHis 471
Db 1252 TGTGTC-----CCGGCTTGCCCTACAACCTTCACCATCCATCCGCGCTGCCCCCTAGA--- 1302
QY 472 TyrAspPheArgGlnSerCys-----TrpValArgAlaIleSerGlnGluAspGln 488
Db 1303 -----TGCCCCCTCCCTGCTCCAAAGCCAAGCCCAAGCCCAACCA 1344
QY 489 LeuAlaProCysTrpGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAla 508
Db 1345 GCAGGCCCAAGAGAGACCGGAA--ACAGATGAAGGGAGAGGAGCAGCTGGAGAGCC 1401
QY 509 ProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGluGlnPro----- 526
Db 1402 CCCAGCCCCCAATCAAGCCATGTGTGTGGCGGCTGCAGAGAGATGATGTCCCCCAGCA 1461
QY 527 CysAlaPro 529
Db 1462 TGTGCCCC 1470

RESULT 5
US-09-774-528-336
; Sequence 336, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)

US-09-774-528-336
Alignment Scores:
Pred. No.: 1.2e-13 Length: 1823
Score: 337.00 Matches: 157
Percent Similarity: 36.2% Conservative: 68
Best Local Similarity: 25.3% Mismatches: 231
Query Match: 11.3% Indels: 167
DB: 3 Gaps: 29

US-10-071-838-2 (1-549) x US-09-774-528-336 (1-1823)

QY 16 GluAspIleIleMetLysTyRGlU--LysGlyHisArgAlaGlyLeuProGluAsp--- 33
Db 154 GAGGACCTGTGCAGCCTCCCGAGCTGCAGATGACTCCAGCTCCTTGGGTCCGACTCA 213
QY 34 -----LysGlyProLysProPheArgSerTyRAsnAsnValAspHisLeuGlyIle 51
Db 214 GAGCTCAGCGGGCTTGCCCATATCGC-----CAGCGGACCGCTATGATTC 261
QY 52 ValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLysGlnIleArgArgGlu 71
Db 262 ATGGGGGAGCTCAGCAGAGCCAGGGCCGGCCACCACCTGCAGACCTCATCGCCAA 321
QY 72 IleSerArgLysSerLysTrpValAspMetLeuGlyAspTrpGluLysTyRlysSerSer 91
Db 322 -----CGGAGATGAAGTGGTGAATGATGACCTCGCATGGGAGAAAACCATGTCCGG 375
QY 92 Arg-----LysLeuIleAspArgAlaTyRlysGlyMetProMetAsnIleArgGlyPro 109
Db 376 CGGTACAAGAAGTAAGATGCAGTCCGGAAGGCATCCCGTCTGCCCGCGCCGA 435
QY 110 MetTrpSerValLeuLeuAsnIleGluMetLysLeuLysAsnProGlyArgTyRgln 129
Db 436 TGCTGCCCCCTGTGTGTGGGGCCCATGTGTGCCAAGAACAAGCCCTGGCAGCTATCAG 495
QY 130 IleMetLysGluLysGlyLysArgSerSerGluHisIleGlnArgIleAspArgAspVal 149
Db 496 GAGCTGGCAGAG--GCCCTGAGAACCCACAGTGAAGAGAACCATTTGGCAGGAGCCTG 552
QY 150 SerGlyThrLeuArgLysHisIlePhePheArgAspArgTyRgLyThrLysGlnArgGlu 169
Db 553 CACCGTCAATTCCTCTGCACGAGATGTTGTGTGCGCTCAGGGCCACGGGACAGGGG 612
QY 170 LeuLeuHisIleLeuLeuAlaTyRglnGluTyRAsnProGluValGlyTyRcysArgAsp 189
Db 613 CTCCTGCAGGTGCTCAAGGCTTACACCTGTATGCACCGGAGCAGGGCTACTGCCAGGCC 672
QY 190 LeuSerHisIleAlaLeuPheLeuLeuTyRLeuProGluGluAspAlaPheTrpAla 209
Db 673 CAGGGGCCCGTGGCTGTGTGCTGCTCATGCACCTGCCCCAGAGAGAGCCTTGTGTGC 732
QY 210 LeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGly 229
Db 733 CTGTGCAATCTGT-----GAGGTCTACCTCCCTGGGTACTACGGGCCCAAC--- 780
QY 230 GlyThrValGlnGlyLeuGlnAspGlnGlnGluHisValAlaIleThrSerGlnProLys 249
Db 781 -----ATGAGGCTGTGCGGCTGAGCGGAGGTGTTCATGAGCCCTGCTGCGG---CGG 831
QY 250 ThrMetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeu 269
Db 832 CTGCTTCGCACTGCACAAAGACCTGCAGAGTGGCGGTGGACCCCTGTACTG 891
QY 270 -----IleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeu 285
Db 892 CCCGAGTGTCTCTGTGCTCTTGGCCGCTCCCTGCCCTTCCACAGTGTGCTGTGTC 951
QY 286 TrpAspValTyRLeuValGluGlyGlnAlaLeuMetProIleThrArgIleAlaPhe 305
Db 952 TGGATGCTTCTCTCAGTGAAGGTGC-CAAGTACTGTTCG----- 992
QY 306 LysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCys 325

```
Db 993 -----TGTGG----- 998
QY 326 AsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuYHisLeuArgAla 345
Db 999 -----GCTGACACTGTGTGCGCTGGCGCTGG-----CACTGCAGAGCA 1037
QY 346 SerMetLysLeuThrArgLysGlnGlyAspLeuPro---ProProAlaLysProGlu 364
Db 1038 GCGAGG-----GCGCTGCCCTGCTCTTGA-----GAC 1067
QY 365 GlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCybLys 384
Db 1068 ACTGGAGCCCTTCGAGCCATCCCCCGCGCAGCTGCAGAGAGAGCCCTCATGTACA 1127
QY 385 GlyAspArg-----GlnAlaProProGly 392
Db 1128 GGTGCACAGCGTGTGCTGTACAGCGGAGCTGCAGCGGAGATCAAGCCCAAGCTGGC 1187
QY 393 ProProAlaArgPhe-----ProArgProIleTrpSerAlaSerProProArg 408
Db 1188 CCAGCTGCCCGATTCCGCGCGGAGCCCGCGCCG-----GCCACAGGTCCGCTCGC 1241
QY 409 -----AlaProArgSerSerThr----- 414
Db 1242 CGGGCCCAAGCCATCTTTGAGGCCCAAGCAGCTGCAGAGTGCAGAGCGCAAGCC 1301
QY 415 -----ProCybProGlyGlyAla----- 420
Db 1302 TGAGTGCCCTCGGATTGTGTGTGACAGCCCGGAGAGAGCCCAAGCCCGCGGAAACC 1361
QY 421 -----ValArgGluAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeu 438
Db 1362 CCAGACCCCGCGCAAGAC---TTTCCATGGGCTCCTGACTCGGCGCGCGGCCCCC--- 1415
QY 439 AlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeu 458
Db 1416 CATCGAGGGGCCCCCAGGCCCCCAAGAGG-----CTCCACCTCTT 1457
QY 459 ProThrAspLeuAspValGlyProTrp----- 468
Db 1458 CCTGACACCCGCTTGTGAGAGACCACTTAGTGTCCCACTCAATGCTGA 1517
QY 469 -----PheArgHisTyrAspPheArgGlnSerCybTrp 479
Db 1518 TGGCTGATGCCAGCCCGCAATAGGCAAGCACTTACTCTTGGGACTCGGGAATTGG 1577
QY 480 ValArgAlaIleSerGlnGluAspGlnLeuAlaProCybTrpGlnAlaGluHisProAla 499
Db 1578 CTT-----CCTTCCTGGCAAGACCAAGCGCAGTGGG 1607
QY 500 GluArgValArgSerAlaPheAlaAlaProSerThrAspSerAsp----- 514
Db 1608 GAAAGAGG-AGGTCCTCGTGTACATAGTGGGTCAAGCACTAGCATGAGAGGGGTCA 1666
QY 515 -----GlnGlyThrProPheArgAlaArgAspGlnProCyb 527
Db 1667 AGAGTGGGCACTGAGGACCCATGAACCTGCTGTGCGCAAGCCCTCACAA----- 1720
QY 528 AlaProThrSerGlyProCybLeuCybGlyLeuHisLeuGlnSerSerGlnPheProPro 547
Db 1721 GTACCAAAAGCCAGCAACAAAGAGTCAAGGGAAGGGTGGCTGAGTCAAGGAGCA 1780
QY 548 Gly 548
Db 1781 GGG 1783
```

```
RESULT 6
US-10-120-988-336
; Sequence 336, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
```

```
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Duntui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6919193el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; PRIOR APPLICATION NUMBER: 2002-04-11
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
US-10-120-988-336
```

```
Alignment Scores:
Pred. No.: 1.2e-13 Length: 1823
Score: 337.00 Matches: 157
Percent Similarity: 36.2% Conservative: 68
Best Local Similarity: 25.3% Mismatches: 231
Query Match: 11.3% Indels: 167
DB: 3 Gaps: 29
```

US-10-071-838-2 (1-549) x US-10-120-988-336 (1-1823)

```
QY 16 GluAspIleIleMetLysTyrGlu---LysGlnHisArgAlaGlyLeuProGluAsp--- 33
Db 154 GAGGACCTGTGTGACGCTCCCGAGCTGCAGAGTACTCCAGCTCTTGGGTCGACTCA 213
QY 34 -----LysGlyProLysProPheArgSerTyrAsnAsnValAspHisLeuGlyIle 51
Db 214 GAGCTCAGCGGCGCTGCGCCCATATCGC-----CAGCGCAGCCGTATGATTC 261
QY 52 ValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLysGlnIleArgGlu 71
Db 262 ATTGGGGGCACTCAGCAGAGCCAGGCGCGGCCACCCACCTGCAGACCTCATCGGCCAA 321
QY 72 IleSerArgLysSerLysTrpValAspMetLeuGlyAspTrpGluLysTyrLysSerSer 91
Db 322 -----CGGAGATGAAGTGGGTGAGATGACCTCGCACTGGGAGAAACCATGTCTCCG 375
QY 92 Arg-----LysLeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyPro 109
Db 376 CGGTACAGAAAGTAAAGATGACGTGCGGAAAGGCATCCCTGTCCCTGCGCGCCGA 435
QY 110 MetTrpSerValLeuLeuAsnIleGluGluMetLysLeuLysAsnProGlyArgTyrGln 129
Db 436 TGCTGCGCCCTGTGTGTGGGGCCATGTGTCCAGAAACACACCCCTGGCACCTATCA 495
QY 130 IleMetLysGluLysGlyLysArgSerSerGlnHisIleGlnArgIleAspArgAspVal 149
Db 496 GAGTGGCAGAG--GCCCTGAGAGCCACAGATGAGAGACCATTTGCAAGGAGCTG 552
QY 150 SerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGlu 169
Db 553 CACCGTCAATTCCCTGTGCAGAGATGTTGTGTGCGCTCAGGGGCCACGGCAGAGGGG 612
QY 170 LeuLeuHisIleLeuLeuAlaTyrGluGluTyrAsnProGluValGlyTyrCybArgAsp 189
Db 613 CTCCTGCAGGTGCTCAAGGCGCTACACCTGTATGACCCGAGCAGGGCTACTGCGAGGCC 672
QY 190 LeuSerHisIleAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPheTrpAla 209
Db 673 CAGGGGCGCGTGGCTGTGTGTCTCATGCACTGCCCCCAGAGGAGGCGCTTCTGTGTC 732
```

QY 210 LeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGly 229
Db 733 CTGGTCAGATCTGT-----GAGGTCTACCTCCCTGGGTACTACGGGCCACAC--- 780
QY 230 GlyThrValGlnGlyLeuGlnAspGlnGlnGlnHisValAlaIleThrSerGlnProLys 249
Db 781 -----ATGAGAGCTGTGCGGCTGAGACGCCGAGGTGTTCATGGCCCTGCTGCGG---CGG 831
QY 250 ThrMetGlyHisGlnAspLysLysAspLeuGlyGlnCysSerProLeuGlyCysLeu 269
Db 832 CTGCTTCGACCGTCACAAAGCACCCTGCAGCAGGTGGCGCTCGAACCCCTGCTGACCTG 891
QY 270 -----IleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeu 285
Db 892 CCCGAGTGGTTCCTGTGCTCTTCGCCCGCTCCCTGCCCCACAGTGCCTGCGTGC 951
QY 286 TrpAspValTyrLeuValGlnGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPhe 305
Db 952 TGGGATGCCTTCTCAGTGAAGGTGC-CAGAGTACTGTTCCG----- 992
QY 306 LysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCys 325
Db 993 -----TGTGG----- 998
QY 326 AsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAla 345
Db 999 -----GCTGACACTGGTGCCTGCGCTGGG-----CACTGCAGAGCA 1037
QY 346 SerMetLysLysLeuThrArgLysGlnGlyAspLeuPro---ProProAlaLysProGlu 364
Db 1038 GCGAGG-----GGCCTGCCCTGCGCTCTGGA-----GAC 1067
QY 365 GlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCysLys 384
Db 1068 ACTGGAGCCCTTCAGACCATCCCCCGCGCAGCTGCAGAGAGAGCCTTCATGTCACA 1127
QY 385 GlyAspArg-----GlnAlaProProGly 392
Db 1128 GGTGCACAGCGTGTGCTGTCAAGCGGACCTGCAGCGGAGATCAAGGCCACAGCTGGC 1187
QY 393 ProProAlaArgPhe-----ProArgProIleTrpSerAlaSerProProArg 408
Db 1188 CCAGCTGCCCGATTCGCGCGCGGACCCCGCCCG-----GCCACAGTCCGCTGCG 1241
QY 409 -----AlaProArgSerSerThr----- 414
Db 1242 CGGGGCCCAAGCCATCTTGAGGCCCAAGCAGCTGGCAGAGTGCAGAGGCCCAAGCC 1301
QY 415 -----ProCysProGlyGlyAla----- 420
Db 1302 TGAGGTGCTCGGATTTGTGTGTCAGACCCCGGAGAGGCCAGACCAACCGCGGAACC 1361
QY 421 -----ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeu 438
Db 1362 CCAGACCCCGCGCAAGAC--TTTCCATGGGCTCTGACTCGGCGCGGGCCCCC--- 1415
QY 439 AlaGlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeu 458
Db 1416 CATCGAGGGCCCCCAGGCCCAACGAGG-----CTCACCTCTT 1457
QY 459 ProThrAspLeuAspValGlnGlyProTyr----- 468
Db 1458 CCTGACACCCGCTTGTAGAGGACCATGACTTAGTCCCCCAGTCTCAATTGCTGA 1517
QY 469 -----PheArgHisTyrAspPheArgGlnSerCysTyr 479
Db 1518 TGGCTGATGCCAGCCCGCAATAAGCACCCTTACTCTTGGGACTCGGGGACTTGG 1577
QY 480 ValArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnHisProAla 499
Db 1578 CTT-----CCTTCTCGCAGAGGACCGAGGACGAGCTGGG 1607
QY 500 GluArgValArgSerAlaPheAlaAlaProSerThrAspSerAsp----- 514

Db 1608 GAAGAGG-AGGTCCTCCGTGTACTACTGGGTCAAGGCACTAGCATGAGAGGGTCAAC 1666
QY 515 -----GlnGlyThrProPheArgAlaArgAspGlnProCys 527
Db 1667 AGAGTGGGGCAGCTGAGAGACCCCATGGAACCGTCTGTCGCCAGGCCCTCACAA----- 1720
QY 528 AlaProThrSerGlyProCysLysLeuHisLeuGlnSerSerGlnPheProPro 547
Db 1721 GTACCAAGCCAGCACCAAGAGTCAAGGAGGGGTGGCTGAGTCAAGGAGCCACAGA 1780
QY 548 Gly 548
Db 1781 GGG 1783

RESULT 7
US-09-949-016-2139
; Sequence 2139, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2139
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2139

Alignment Scores:
Pred. No.: 2.59e-13 Length: 1925
Score: 332.50 Matches: 119
Percent Similarity: 39.6% Conservative: 57
Best Local Similarity: 26.8% Mismatches: 135
Query Match: 11.1% Indels: 133
DB: 3 Gaps: 15

US-10-071-838-2 (1-549) x US-09-949-016-2139 (1-1925)
QY 113 ValLeuLeuAsnIleGlnGluMetLysLeuLysAsnProGlyArgTyr---GlnIleMet 131
Db 30 GTAGCACTCAAACTTGCCACGAGCGGAGCTGAATAAGTCTAAATATGACAGAAATTA 89
QY 132 LysGluLysGlyLysArgSerSerGlnHisIleGlnArgIleAspArgAspValSerGly 151
Db 90 AAACACAGACACGGGCTGTTCACTGACATCAGACAAATAGACCTGGATGTCACCGC 149
QY 152 ThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeu 171
Db 150 ACATTTCGGGACCAATTATGTTTAAAGACAGATATGTTTAAACAAATCTTATTTC 209
QY 172 HisIleLeuLeuAlaTyrGlnGluTyrAsnProGluValGlyTyrCysArgAspLeuSer 191
Db 210 CATGTGCTGCTGCTATTTCTATTATTAACACGGAATCGGGTATTGTCAAGGATGAGC 269
QY 192 HisIleAlaAlaLeuPheLeuLeuTyrLeuProGlnGluAspAlaPheTrpAlaLeuVal 211
Db 270 CAGATCACAGCTTTACTCTCATGTATATGAACGAGGAGATGCCCTTCGGGCCCTGCTC 329
QY 212 GlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThr 231
Db 330 AAACCTTCTCAGGCGCCTAAACATGCCATG----- 359

```

QY      232 ValGInGlyLeuGInAspGInGInGluHISValAlaThrSerGInProLysThrMet 251
Db      360 -----CATGAACATCTAATGAATAATTGTCATGGAGAAGAA----- 392
QY      252 GlyHISGInAspLysLysAspLeuCyseGlyGInCyseSerProLeuGlyCyseLeuIleArg 271
Db      392 ----- 392
QY      272 IleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuVal 291
Db      393 -----CTTGTA 398
QY      292 GluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGInGInLysArg 311
Db      399 GAA----- 401
QY      312 LeuThrLysThrSerArgCyseGlyProTrpAlaArgPheCyseAsnArgPheValAspThr 331
Db      402 -----TTTTTCAGGAGACC 416
QY      332 TrpAlaArgAsp-----GluAspThrValLeuLysHISLeuArgAlaSerMet 347
Db      417 CTGGCAAAAGATTTTTCTTTGAAAGATGATTTTGTGATAGAGCAACTTCAGATTTCATG 476
QY      348 LysLysLeuThrArgLysGInGlyAspLeuProProProAlaLys-----Pro 363
Db      477 ACAGAACTAAAGCGGCAAAAGTTAGACTCTCCAGAACCTGGTAAAGAGATGAATATCCA 536
QY      364 GluGInGlySerSerAlaSerArgProValProAlaSerArgGlyLysThrLeuCyse 383
Db      537 AAGAAGCCCTTGGGGCAGCTTCCACCTGAATTCTAGTCTTGGGGCTCCATCATTGAGC 596
QY      384 LysGlyAspArgGlnAlaProProGlyProPro-AlaArgPhePro----- 398
Db      597 AACGACACAGAGAGCGCTG-----GGCCGGCCGAGCCCGCTGGCCAGCGGACAGAGGAG 650
QY      399 -----ArgProIleTrpSerAlaSerProArgAlaProArgSerSerThrProCy 416
Db      651 AGCGGGGCGCCCAAGAGAGGACGAGCACTCCCGCACCCCGCAGACAGACCGGAGC 710
QY      416 sProGlyGlyAlaValArgGluAsp-----ThrTyrProValGlyThrGInGlyValPr 434
Db      711 CCCGAGAGAGCACAGCCCGCAAGACGAAATCGGTGAGAGAGAGAGCAAAAGCTTAA 770
QY      434 oSerProAlaLeuAlaGInGlyGlyProGInGlySerTrpArgPheLeuGInTrpAsnSe 454
Db      771 GATGAGGACAGATTTCAAAGAAAACGCCATCGGGTCCACACAGACAGTCCAGGCAATAT 830
QY      454 rMetProArgLeuProThrAspLeu-AspValGluGlyProTrpPheArgHISTrpAsp 474
Db      831 AATCACGCGAGCTGCCAACCMAAATAGCAA-----CGCCACTTCAAT 872
QY      474 heArgGInSer---Cys-----TrpValArgAlaIleSerGInGluAspG 488
Db      873 ATCAGGAAGAGAGTTGTGTGCCCAATGAATAAACCGTCAGACGTCCTACGTACAGAGAGA 932
QY      488 InLeuAlaProCystrpGlnAlaGluHISProAlaGluArgValArgSerAlaPheAla 508
Db      933 AC-----TGC---CAATACACCATGGAAGGCMAAGTGCAGACGACGCCCGCGC 980
QY      508 Ia-----ProSerThrAspSerAspGInGlyThrProPheArgAlaArgAspG 524
Db      981 CTGCGAGTTACCGTCCACAGTCTGCGGAGGTGCGGTGTCAAAACGTGCGGCCAAGATG 1040
QY      524 IuGInPro 526
Db      1041 AAGGCCCT 1048

```

RESULT 8
 US-09-620-312D-675
 ; Sequence 675, Application US/09620312D
 ; Patent No. 6569662

```

; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 675
; LENGTH: 3039
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (169)..(2502)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(3039)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-675

Alignment Scores:
Pred. No.: 4,97e-08 Length: 3039
Score: 258.00 Matches: 157
Percent Similarity: 35.9% Conservative: 77
Best Local Similarity: 24.1% Mismatches: 236
Query Match: 8.6% Indels: 183
DB: 3 Gaps: 30

US-10-071-838-2 (1-549) x US-09-620-312D-675 (1-3039)
QY      9 SerTrpAlaGInGInuArgGluAspIleIleMetLysTyr---GluLysGlyHISArg 27
Db      220 AGCATATGCGCCCAAGAG-----ATCTTGGCCAAGTACACGACGAGAAGAGACTCA 270
QY      28 AlaGlyLeuProGlu-----AspLysGly 35
Db      271 GCAGAGCAACAGAGTTCTACTACGATGAGTTGTGTTCCGTGTACAAAGAAAGAGGT 330
QY      36 ProLysProPheArgSerTyrAsnAsnAsnValAsp----- 47
Db      331 GATGAGCCTGGCTCCAGTCTGTGCGGAACCTCCCTGTATGAGAGATGCTCCACAGAGG 390
QY      48 -----HISLeuGlyIleValHISGluThrGluLeuProProLeuThrAla 62
Db      391 CTGCGGTGCGACAGCCCACTGTGAGTTCACCCATTAACACGATGTGGGAGATCTCAC-- 447
QY      63 ArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAspMetLeu 82
Db      447 ----- 447
QY      83 GlyAspTrpGluLys-----TyrLysSerSerArgLysLeuIleAspArgAla 98

```

```
Db 448 -----TGGAACAAGATTGCCGTCTCCCTACCCCGCTGTGAGAAGCTCCGCTCCCTGGTG 501
      |||::|||
Qy 99 TyrLysGlyMetProMetAsnIleArgGlyProMetTrpSerValIleuAsnIleGlu 118
      |||::|||
Db 502 CTGGCCGGCATCCCAATGGCATGAGGCCACAGCTGTGATGCGGCTCTCTGGGGCCCTG 561
      |||::|||
Qy 119 GluMetLysLeuLysAsnProGlyArgTyr---GlnIleMetLysGluLysGlyLysArg 137
      ::::|
Db 562 CAGAAGAAGAGAACTCTGAGCTGTCTTACCCGGAAGATTGTGAAGAAACAGCTCCAAAGAT 621
      ::::|
Qy 138 SerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIle 157
      ::::|
Db 622 GAGACCATCGCTGCCAAGCAGATCGAAGAGAACTGTCTCCGACCAATGCCAGCAACGCC 681
      ::::|
Qy 158 PhePheArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyr 177
      |||
Db 682 TGCTTCGCCAGCATGGGTAGCATCGGGTGCCTGCGCAGGGTGTCCGGCCCTG 741
      |||
Qy 178 GluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPhe 197
      |||
Db 742 GCCTGGCTTACCCAGAGATCGGCTACTGCCAGGGCAACGCGCATGGTGGCCGCTGCTC 801
      |||
Qy 198 LeuLeuTyrLeuProGluGluAspAlaPheTrp-----AlaLeuValGln---Leu 213
      |||
Db 802 CTGCTGTCTCTGAGAGAGAGACGCGCTTCTGAGATGATGTCTGCCATCATCGAAGAACCTG 861
      |||
Qy 214 LeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGln 233
      |||
Db 862 CTCCCGCGCTCTACTTCAACACC-----ACCCTGCTG 894
      |||
Qy 234 GlyLeuGlnAspGlnGln-----GluHisValValAlaThrSerGlnProLysThr 250
      |||
Db 895 GGTGTCCAGACTGACCAAGCGGGTCTGCGCCCACTTGTTCAGTAACCTGCTGCGCTG 954
      |||
Qy 251 MetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIle 270
      ::::|
Db 955 GACAAGCTGCTCCAGAGCATGACATGAGCTGTCCCTGATCACACTGCACTGGTTCTC 1014
      |||
Qy 271 ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeu 290
      |||
Db 1015 ACGGCCTTCGCCAGCGGTGGTGGACATCAAGCTGCTCCTGCGCATCTGGGACCTGTTTTC 1074
      |||
Qy 291 ValGluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLys 310
      |||
Db 1075 TACGAGGGCTCCCGGGTGTGTTCAGACTCAGCTGGGCACTGCTGCAAGGAGAA 1134
      |||
Qy 311 ArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAsp 330
      |||
Db 1135 GAGCTGATCCAGTCAAGAACTCG-----GCCTCATCTTCAACACGCTATCGGAT 1185
      |||
Qy 331 ThrTrpAlaArgAspGluAspThr-----ValLeuLysHisLeuArgAla 345
      ::::|
Db 1186 ATCCCGTCGACAGATGAGAGACGCGAGCTGTCTGGGGTGGCCATGCGGCTGCGCGC 1245
      |||
Qy 346 SerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro----- 359
      |||
Db 1246 TCCCTCAC-CGATGTGGCGT-----GGAAGTCAAGCGCCGCAAGCTCTGCGCTATCT 1298
      |||
Qy 360 -----ProAlaLysProGluGlnGlySer 367
      |||
Db 1299 CATTGCAGACCAAGGCGAGCTCTGGGGCGCGCACTTCACCAACTCTCTCAGGTTGT 1358
      |||
Qy 368 SerAlaSerArgProVal-----ProAlaSerArgGlyGly--- 379
      |||
Db 1359 TCGCCGAGAGACCAAGGAGAAAGTCCACCATCACTGCTGTCTTCGGGAGAGATGA 1418
      |||
Qy 380 -----LysThrLeuCysLysGlyAspArgGlnAlaProProGly 392
      ::::|
Db 1419 CCTGAGGCACTCAAGGCCAAGAACATCAAGACGAACTGTGTGCTGACCTCCGGGA 1478
      |||
Qy 393 Pro---ProAlaArg---PheProArgProIleTrpSerAlaSerProProArgAlaPro 410
      |||
```

```
Db 1479 AGCCATCTGCGCGGTGGACCGCCCACTTCCAGTG-----CACAGACCCCAAAACTGCAG 1532
      |||
Qy 411 ArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTyrProValGlyThr 430
      |||
Db 1533 CGTGTGAGTGCACAGCTCCCTGGCTGT-----ACC 1565
      |||
Qy 431 GlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGln-GlySerTrpArgPheLe 450
      ::::|
Db 1566 AAACACGGCCCTAACTCCTCAACCCCTTGTGTGGCTGTGTCTCACTGTGCAGAGCT 1625
      ::::|
Qy 450 uGlnTrpAsn---SerMetProArgLeuProThrAsp----- 461
      ::::|
Db 1626 GACTCCAGACTATAAGCATGAGAGCCACAGCGGACCAAGAACTACGTGGCGTGTCTC 1685
      ::::|
Qy 462 -----LeuAspValGluGlyProTrpPheArgHis 471
      |||
Db 1686 ACGCAGCCACCGCGCGCGAGCCCAAGCCCTGCTGAGCTTTGAG-----CGGCA 1733
      |||
Qy 471 sTyrAsp-----PheArgGlnSerCysTrpValArgAlaIleSerGlnGluAs 487
      |||
Db 1734 CGACGACGACGAGCTGGGCTTCCGAAAGAACACATCATCAATCGTGTCTCAGAAGGA 1793
      |||
Qy 487 pGlnLeuAlaProCysTrpGlnAlaGlu-----HisProAlaGln 500
      |||
Db 1794 CGAG-----CACTGTGGTGGGAGCTCAACGCGCTGCGAGGCTGTTTCCAGCCAA 1847
      |||
Qy 500 uArgVal-ArgSer---AlaPheAlaAlaProSerThrAspSerAspGlnGlyThrPro 519
      |||
Db 1848 GTTCGTGAAGTCTCTGATGAGCGCGCAAGAGTACTCCATCGCGGGAGATGATCGGT 1907
      |||
Qy 519 heArg-----AlaArgAspGluGlnProCysAs 528
      |||
Db 1908 GACGAGGGGTCAACAGACTCGTGCGAGGAGACCTCTGCGCGGCCCTTAAGCCCTGTT 1967
      |||
Qy 528 laPro-----ThrSerGlyProCysLeu 535
      ::::|
Db 1968 CGAACATGGACTGAAGAAAGCCATCCCTGCTTG 1999
      |||
```

```
RESULT 9
US-09-620-312D-676
; Sequence 676, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 676
; LENGTH: 2955
```



```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (169)..(2418)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2955)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-676

Alignment Scores:
Pred. No.:      1.04e-07
Score:          253.00
Percent Similarity: 35.8%
Best Local Similarity: 24.6%
Query Match:    8.5%
DB:             3

```

Alignment Scores:

Pred. No.:	1.04e-07	Length:	2955
Score:	253.00	Matches:	156
Percent Similarity:	35.8%	Conservative:	71
Best Local Similarity:	24.6%	Mismatch:	233
Query Match:	8.5%	Indels:	175
DB:	3	Gaps:	28

US-10-071-838-2 (1-549) x US-09-620-312D-676 (1-2955)

[illegible]

QY	234	GLYLEUGLNASPGINGLN-----GLUHSIVALVALAAThrSerglnProlySthr	250
		:::	
Db	895	GGTGTCCAGACTGACCAGCGGGTCTGCGCCACCCTATTGTCCAGTACCTGCGCTG	954
QY	251	MetGlyHisGlnAspLysLysAspLeuCySgIyInCySerProLeuGlyCySLeuIle	270
		:::	
Db	955	GACAAGCTGCTCCAGAGCATGACATTTAGCTGTCCCTGATCACACTGCACCTGTTCTC	1014
QY	271	ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrlu	290
		:::	
Db	1015	ACGGCCTTCGCCACGCGTGTGGACATCAAGCTGCTCCTGCCATCTGGAGCCTGTTTTC	1074
QY	291	ValGluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLys	310
		:::	
Db	1075	TACGAGGGCTCCCGGGTGTCTTCCAGCTCACGCTGGGCATGCTGCACCTCAAGAGGAA	1134
QY	311	ArgLeuThrLysThrSerArgCySgIyProTrpAlaArgPheCySAsnArgPheValAsp	330
		:::	
Db	1135	GAGCTGATCCAGTCAGAGAACTCG-----GCCTTCATCTTCAACACGCTATCGGAT	1185
QY	331	ThrTrpAlaArgAspGluAspThr-----ValleuLysHisLeuArgAla	345
		:::	
Db	1186	ATCCCGTCGACATGAGAGACGCGGAGCTGCTTCTGGGGGTGGCCATGCGGCTGCCGCC	1245
QY	346	SerMetLysLysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGluGln	365
		:::	
Db	1246	TCCCTCAC-CGATGTGGCGCT-----GGAGACTCAGCGCCGCAAGCT-----CCT	1289
QY	366	GlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCySlySgIy	385
		:::	
Db	1290	GGCCTATCTCATTTGCAGACCGGGCCAGCT-----	1319
QY	386	AspArgGlnAlaProProGlyProProAlaArgPheProArgProIleTrpSerAlaSer	405
		:::	
Db	1320	-----CTGGGGGCGCGCACCTTCACCAACCTTCTCAGGTTGT	1358
QY	406	ProProArgAlaProArgSerSerThrProCySProGlyGlyAlaValArgLysAspThr	425
		:::	
Db	1359	TCGCGCAGAGACCAGCGGAGAAATCCACCATTACTGCTCTGCTCTCGG---GAGGA	1415
QY	426	TyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyGly---ProGln	444
		:::	
Db	1416	TGACCTGAGGACACTCAAGGCCAAGAATCAAGCAGACGGAATGATGGCTGACCTCCG	1475
QY	445	GlySerTrp-----ArgPheLeuGlnTrp	452
		:::	
Db	1476	GGAAGCCATCTCGCGCGTGGCACGCCACTTCCAGTGCACAGACCCCAAAACTGCAGCGT	1535
QY	453	AsnSerMetProArgLeu-----ProThr-Asp-----	461
		:::	
Db	1536	GGAAGTACTCCAGACTATAGCATGAGAGCCACCGGGGACCAAGAACTACGTGGC	1595
QY	462	-----LeuAspValGluGlyProTrpPh	469
		:::	
Db	1596	GTCGTACGACGACCCAGCGCGCGAGCCAAAGCCCTGTGACTTTGAG-----	1644
QY	469	EAArgHisTyrAsp-----PheArgGlnSerCySTrpValArgAlaIleSergI	485
		:::	
Db	1645	-CGGACGACGACGACGAGCTGGGCTTCCGCAAGAACATCATCAAAATCGTGTCTCA	1703
QY	485	nGluAspGlnLeuAlaProCySTrpGlnAlaGlu-----HisPr	498
		:::	
Db	1704	GAAAGACGAG-----CACTGCTGGGTGGGGAGCTCAAGGCTGCGAGGCTGTTCC	1757
QY	498	oAlaGluArgVal-ArgSer--AlaPheAlaAlaProSerThrAspSerAspGlnGlyT	517
		:::	
Db	1758	AGCCCAAGTTCGTGAAGTCTTGATGAGCGCAGCAAGAAGTACTCCATCCGGGGGATGA	1817
QY	517	hPrPheArg-----AlaArgAspGluGlnP	526
		:::	
Db	1818	CTCGGTGACGAGGGGGTCAAGACCTCGTGCAGAGGACCTCTGCGCCGCTTAAGGC	1877
QY	526	roCyAlaPro-----ThrSergIyProCySleu	535


```
Qy 95 ILeAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyProMetTrpSerValIleu 114
Db 492 AAGGAATTGTTCTAAAGGGAATACCCATCACTTTAGAGCAATAGTTGGCACTTTTA 551
Qy 115 LeuAenIleGluGluMetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLys 134
Db 552 TGCAGTGCACAAGATATGCCAATTAAAGATCAG-----TATTGAGAACTC 596
Qy 135 GlyLysArgSerSerGlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArg 154
Db 597 CTGAAATATGACCTCGCCTTGTAATAATTGATCCGAAGGACATTGCTAGAACTTAACCT 656
Qy 155 LysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeuLysIleLeu 174
Db 657 GAACACAACCTTTTAAAGAAAAGATAGCCTTGACAGAGGCTTTATTATTAAATGTAATG 716
Qy 175 LeuAlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAla 194
Db 717 AAGGCTTACTCTTATAGATCGTAGAGGTGTTACTGTCAAGAGAGCTTTTATTATGTAATG 776
Qy 195 AlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeu 214
Db 777 GGAATGTTGCTTATGACAGATGCCAGAAGAAGCTTCTGTATATTGTTAAATTAATG 836
Qy 215 AlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly 234
Db 837 CAAGATTATAGA-----CTTCGTGAACCTTTTAAACCAAGTATGCGAATTTGGCCCTT 890
Qy 235 LeuGlnAspGlnGlnGlnHisValAlaIleThrSerGlnProLysThrMetGlyHisGln 254
Db 891 TGTATGTATCCAGTTGAAATGATGATACAGAGGACATCTTCCAGAGCTCTTGTACATTTT 950
Qy 255 AspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIle 274
Db 951 CAATCTCAGAGTTTCAATACCTCAATGTATGATCATCATTCCTGTTCTGACTATCTTCTT 1010
Qy 275 AspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlu 294
Db 1011 ACAACTTTTCCACTACCAAGTCCAAACAAGATATTGATATCTTATGTCTGAGGGTTTA 1070
Qy 295 GlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnIleLysArgLeuThrLys 314
Db 1071 GAATAGTGTTCGTGTAGAGTTAGCACTTCTTCAGATGAATCAGGCAAGCTGATGCAA 1130
Qy 315 ThrSerArgCysGlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArg 334
Db 1131 CTTCACATGGAAGGATGTATACAGCACTTCAAAAGCTCATTCACATCAGTTTGATGCT 1190
Qy 335 AspGluAspThrValLeu-----LysHisLeuArgAlaSerMetLys 348
Db 1191 GTCCAGACAAAGTAAATCCAGCAGCTTACCAGTCAATACAAATTCAAAAAAATGAAA 1250
Qy 349 LysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGluGln 365
Db 1251 AAGCTTGAAAAAGAAATACATAATAAAAAAGAAAGAAATGGAAGAGCAA 1301

RESULT 12
US-09-270-767-12482
; Sequence 12482, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12482
; LENGTH: 1887
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-12482
```

```
Alignment Scores:
Pred. No.: 7.13e-06 Length: 1887
Score: 222.00 Matches: 129
Percent Similarity: 36.6% Conservative: 64
Best Local Similarity: 24.5% Mismatch: 198
Query Match: 7.4% Indels: 138
DB: 3 Gaps: 23

US-10-071-838-2 (1-549) x US-09-270-767-12482 (1-1887)

Qy 47 AspHisLeuGlyIleValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLys 66
Db 105 GACCTTCTGGCTCGAGTTCAAT-----GTTCTTTGAGTCGCGAAGCGCTAAG 152
Qy 67 -----GlnIleArg 69
Db 153 TACGACATTTCTGGAGCAAGCAGACCGCTTGATGAACACATTCAAGACAGATTTCAGC 212
Qy 70 ArgGluIleSerArgLysSerLysTrpValAspMetLeuGlyAspTrpGluLysTyr-- 88
Db 213 GCAGAGATCATACAGAAAGCAGAG-----GAAAGATGCTGCGTGGAGCGGCATTTC 266
Qy 89 -----LysSerSerArgLysLeuIleAspArgAlaTyr 99
Db 267 CCGGACTTTCGACGGGCAATTGGAATGTTCCGCACAAAGCAGCATGATTAATCTAATAGTG 326
Qy 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAenIleGluGlu 119
Db 327 GAGGATATCTCTGACAAACTGCGCCAGAGATCTGGCTGATATTCTCGGGGCAATCCAC 386
Qy 120 MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGly-----Lys 136
Db 387 GACAAAGAGATGAATCCCGACCTGACAGAGATCTCGTAAGAAAGCGGCTTGCAATAAAG 446
Qy 137 ArgSerSerGlnHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHis 156
Db 447 AACTGCTTTCGCCAC---GATGAATGTATCGGACTTGCCGCGTTCGCTGCCAGCAT 503
Qy 157 IlePhePheArg-----AspArgTyrGlyThrLysGlnArgGluLeuLysIleLeu 174
Db 504 CCGGCAATTCAAGACCGATGCGCATAGGCTGCTGCGGAGA-----GTCCTT 551
Qy 175 LeuAlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAla 194
Db 552 CAGGCTTATGCGCTGCGGAATCCACAAGTGGGCTACTGTCAAGGCAATGATTTGTGTCG 611
Qy 195 AlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeu 214
Db 612 TCGGTGTTCTGCTCTTCTGTGATGAGAGAAAGCATTTCTGATGCTGGCCAGCCTGTGC 671
Qy 215 AlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGly 234
Db 672 -----GAGAACTTACTACCGGACTACTACAAGATAAA-----GTGGTGGGC 713
Qy 235 LeuGlnAspGlnGln-----GlnHisValAlaIleThrSerGlnProLysThrMet 251
Db 714 GCCCAATCGATCAGGCTGTGCTTAATGAGCTAGTGAGACGCAATTCCTGATTGAC 773
Qy 252 GlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArg 271
Db 774 GGGCATCTGAGCAGTGGGTGTGATTAAGATGATTTCC--ATTCTGTGTTCTTACC 830
Qy 272 IleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuVal 291
Db 831 ATCTTCATGAGCGGTATCAGCTACGAGAGCTCGCTGCACATCCCTGACTTTTCTTAC 890
Qy 292 GluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnIleLysArg 311
Db 891 GAGGGCGCCAAAGATCATTTTATGATTTCACTGCACATCATTTGAATGGAACAGGACAAA 950
Qy 312 LeuThrLysThrSerArgCysGlyProTrpAlaArg-PheCysAsnArgPheValAspTh 331
```

```
Db      951 CTGTTGATCTGCCAAGATGATGGCGAAGCCATGCTAGTTTTCAGA-----ACTAC 1001
Qy      331 rTTPAlaArgAspGluAspThrValLeu-----
Db      1002 CTGAGAG-----GAGTTTACAATCCAGAGTACCAGGTGCCTCTTACGACGCAAAACGG 1055
Qy      341 -----LysHisLeuArgAlaSerMetLysLysLeuThrAr 352
Db      1056 AAGATGAGCGTAAAGTCCAAACACAAACAGTTCAGACCCCTGATCCATGAGGCATACACA 1115
Qy      352 glySerGlnGlyAspLeuProProProAlaLysProGlnGlnGlySerSerAlaSerArgPr 372
Db      1116 AAAAGATTCGAT-----GAGGAGTATCACACAGCAGCGCATCGAGGAG 1160
Qy      372 ovalProAlaSerArgGlyGly-----
Db      1161 TTACGCAACAGCATCGCCGGTTAAAGATGCGACAGTTGATTCGACAAACGAGAAGACC 1220
Qy      380 -----LysThr 381
Db      1221 ATTGTCAAGCCTTACGTCAGAAATCCGTATTTCAAACCCAGTGAATCGCATATGCTGCTC 1280
Qy      381 rLeuCybLysGlyAspArgGlnAlaProProGlyPro-----
Db      1281 ACTATCATCCGGGAGGAAAGACAGCGCCTTAATCCCTGCAACAGCAGCAGAGAGAGTT 1340
Qy      394 -ProAlaArgPheProArg--ProIleThrSerAlaSerProProArgAlaProArgSe 412
Db      1341 CAATGCCCGTTGTCTGAGACTCCCAAGCTGCTCCCTCAGT---CGACAAAGCAGACCGCATC 1397
Qy      412 rSerThrProCybProGlyGlyAlaValArgGluAspThrTyProValGlyThrGlnG 432
Db      1398 CAAGATGCCGGCGCCTCC-GGAGGC--AGATACGAGGCTTACAGTGTCAAGCTATGA-CG 1452
Qy      432 yValPro-----SerProAlaLeuAlaGlnGlyGlyProGlnGlySerTr 447
Db      1453 TATTCACACACTTTTACGGAAGTCAACCTGGCGTTAAGTGCCTCAGCGTGGACATTG 1512
Qy      447 PArgPheLeu-----GlnTrpAsnSerMetProArgLeuProThrAs 461
Db      1513 GCGAAAAATTGTTTAGCCTTACTGATGAAGAAGGAAGTGTGTCTCGACTTCGACAGC 1572
Qy      461 pLeuAspValGluGlyPro 467
Db      1573 TCATCAATGCATTGGGCTT 1591

RESULT 13
US-09-270-767-9899/C
; Sequence 9899, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9899
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9899
```

```
Alignment Scores:
Pred. No.:      0.000168      Length:      806
Score:          194.50        Matches:      79
Percent Similarity: 41.1%      Conservative: 44
Best Local Similarity: 26.4%      Mismatches: 111
Query Match:     6.5%         Indels:      67
DB:              3           Gaps:         9
```

US-10-071-838-2 (1-549) x US-09-270-767-9899 (1-806)

```
Qy      101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGluMet 120
Db      788 GGCCTTCCAGAGCACTGCCTGAAAAGATCTGGCAGAACTAGCCAAAGCTCGAGGGC--- 732
Qy      121 LysLeuLysAsnProGlyArgTyArgGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db      731 AGAATGGAATATGAACGACAACTACAAGATCTTAATCACCAAGAAACCAATGTGAGACC 672
Qy      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db      671 GTTATCCAG-----CGGACATCCATCCGACTTCCCGCGCACAAATGCTTCAAA 621
Qy      161 AspArgTyArgLysThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyArgGluGluTyr 180
Db      620 GAAATTGGCGGTTCTGGCCAAAGATGCACCTTTTAAAGTGTCCAAAGCGTATGCCGTTCAT 561
Qy      181 AsnProGluValGlyTyrCybArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db      560 GACAGCGAGGTTGATATGTCAAGGCTTAAGTTTCATAGCAGCTAGTCTGCTCCTTCAT 501
Qy      201 LeuProGlnGluAspAlaPheThrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db      500 ATGCTGACGAGGATGCGCTTCTGTCTGTCTGTAGCGCTTATGTACGAC----- 453
Qy      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAsp----- 237
Db      452 -----TATGGGCTCCGTGATCTCTACAAA 429
Qy      238 -----GlnGlnGluHisValAlaThrSer 246
Db      428 GCCCGATTGAGGTCCTTACCTTCTGTCTTACCAACTGGAGCGACTGATCAAGATCAG 369
Qy      247 GlnProLysThrMetGlyHis-GlnAspLysLysAspLeuCybGlyGlnCybSerProLe 266
Db      368 CTGCCCAAGCTGCACGAACACTTCACGCGCTGGCGCATTGAGACGACATGTACGCTCC 309
Qy      266 uGlyCybLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTr 286
Db      308 CAATGC-TTCCCTAACCTTATATACAGCGGATGCCATTTGTTTGTGTTCCACGTGCT 250
Qy      286 pAspValTyrrLeuValGlnGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLys 306
Db      249 GGATGTGTTTCTACTGACGGA-----CTAAGCTGTG----- 219
Qy      306 sValGlnGlnLysArgLeuThrLysThrSerArgCybGlyProTrpAlaArgPheCysAs 326
Db      218 -CTCTTCAGGTGGCGGTACTCTTATCAATCTGTGAATCC----- 177
Qy      326 nArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSe 346
Db      176 -----GACTTGGCGCAACTCGATTGTTGAGGCGCATTTCTAAAGTATTTCCGGGTAA 127
Qy      346 rMetLysLysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGlnGlnG 366
Db      126 GTTCCGAAAAAAGTGCCGCACTCC-----AGTCAGGC-ACGCAAGG 86
Qy      366 ySerSerAlaSer-----ArgProValProAlaSerArgGlyGly 379
Db      85 TGATGAAGCAAGCCTGCCGAACGTAAAGTAAAAAACTGAAGCAGTACGAGAGGA 31
```

```
RESULT 14
US-08-363-300-1
; Sequence 1, Application US/08363300
; Patent No. 5700927
; GENERAL INFORMATION:
; APPLICANT: Zon, Leonard and Richardson, Paul
; TITLE OF INVENTION: Tbc1 Gene and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
```



```

STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,300
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04590/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4039 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 403..3829
US-08-363-300-1

```

```

Alignment Scores:
Pred. No.: 0.0088      Length: 4039
Score: 182.50          Matches: 125
Percent Similarity: 34.6%      Conservative: 67
Best Local Similarity: 22.5%    Mismatches: 227
Query Match: 6.1%             Indels: 139
DB: 2                      Gaps: 18

```

US-10-071-838-2 (1-549) x US-08-363-300-1 (1-4039)

```

QY 13 GlnGluArgGluAspIlele-----MetLysTyrGluLysGluHisArgAlaGly 29
DB 2408 AGAAAGAGGAGAGCTGACGCGAGCTTCGAGAGCTGTGAAAGGCCATCTTGACACA 2467
QY 30 LeuProGluAspLysGly---ProLysProPheArgSerTyrAsnAsnValAspHis 48
DB 2468 GATCGTCCTCGTCAGATGAGAGAGAGATCAGAACTAAGCTAACAAG-----CCT 2515
QY 49 LeuGlyIleValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLysGlnIle 68
DB 2516 CTGAAGACGATTGCTGTAACAACGCCCTCAAGCTTGACTATGAAGAAATCACTCCGTGC 2575
QY 69 ArgArgGluIleSerArgLysSerLysTyrValAspMetLeuGly----- 83
DB 2576 TT-AAAGAGTCACT-----ACAGTGTGGGAAAGATGCTTAGCACTCCAGGAAGATCC 2628
QY 84 -----AspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyr 99
DB 2629 AAAATTAAAGTTGACATGAAAAAGTGCACTCAGCT-----GTTGGG 2670
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGlu 119
DB 2671 CAAGGTGTGCCACGTCACTACCGAGGTGAGATCTGGAATTCTAGCTGAGCAGTTCAC 2730
QY 120 MetLysLeuLysAsnProGlyArgTyrGln-----IleMetLysGluLysGly 135
DB 2731 CTTAACACCCATTCTCTAGTAACAAGCAGCCAAAGACGTGCCCTACAAGAGCTCCTG 2790
QY 136 LysArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLys 155
DB 2791 AAGAACTGACCTCGCAGCAGCAGCCCATTTCTCATGCACTCGGGGGAACCTTTCCACA 2850

```

```

QY 156 HisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeu 175
DB 2851 CATCCATACTTCTCTGCCAGCTTGAGCAGAGCTGACGTCACTTACACACTTCTGAAG 2910
QY 176 AlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLysSerHisIleAlaAla 195
DB 2911 GCCTACTCGCTTCTTGACCCAGAGGTTGGATATCTGCCAAGGTTCAAGCTTTGTGGCAGGC 2970
QY 196 LeuPheLeuLeuTyrLeuProGluGluAspAlaPheThrAlaLeuValGlnLeuLeuAla 215
DB 2971 ATTTGCTTCTTCAATGAGTGAAGAGAGCGCTTCAAGATGCTCAAGTTCTTGATGTTT 3030
QY 216 SerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlnGlyThrValGlnGlyLeu 235
DB 3031 GAC-----ATGGGCTGCGGAACAGTATCGGCCAGACATGATTATTGGCAGATCCAG 3084
QY 236 GlnAspGlnGlnGluHisValAlaAlaThrSerGlnProLysThrMetGlyHisGlnAsp 255
DB 3085 ATGTACCAAGCTGTACAGGCTCTCCACGATTACACAGAGACCTTACAACCACTGGAA 3144
QY 256 LysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAsp 275
DB 3145 GAGCAGAGACTGGCCCCCTACGTACGGCGCTCCCTGTTTCTCACCGTGTCCCTCA 3204
QY 276 GlyIleSerLeuGlyLeuThrLeuArgLeuTyrAspValTyrLeuValGlnGlyGln 295
DB 3205 CAGTTCCCACTCGGCTTGTGACCAAGAGTCTTGATATGATCTTCTTCAGGATCAAGAG 3264
QY 296 AlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThr 315
DB 3265 -----GTCATATTAAAGTAGCTTAAAGTCTTTTG----- 3294
QY 316 SerArgCysGlyProThrAlaArgPheCysAsnArgPheValAspThrThrAlaArgAsp 335
DB 3295 -----GGGAGCCAT 3303
QY 336 GluAspThrValLeuLysHis-----LeuArgAlaSerMetLysLysLeuThrArgLys 353
DB 3304 AAGCCCTGATTCTACAGCATGAGAACCTGGAACCATCGTGACTTCAATAAGAACACA 3363
QY 354 GlnGlyAspLeuProProProAlaLysProGlnGlnLysSerSerAlaSerArgProVal 373
DB 3364 CTCCCAACCTGGG-CCTGAGTGCAGATGAGAGAACCATCACTCAGTCAAGT----- 3410
QY 374 ProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaProProGlyPro 393
DB 3411 -----GTTGAGATGACATCGCCAAAGCACTCCAGGCCTA 3446
QY 394 ProAlaArgPheProArgProIleTyrSerAlaSerProProArgAlaProArgSerSer 413
DB 3447 TGAGGTGAGTACACAGTGT----- 3467
QY 414 ThrProCysProGlyGlyAlaValArgGluAspThrTyrProVal----- 428
DB 3468 -----CCAGGAGAGCT-----TATTGAGTCTCGCCTCTCAGTGA 3503
QY 429 -----GlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGln 444
DB 3504 CAACCAAGAAATGAGAAATTGAGAAACCAACAGCAGCAGTTCGCGCAACAGAACCTTG- 3562
QY 445 GlySerTyrArgPheLeuGlnTyr----- 452
DB 3563 ACCTCCTGAGCAGTTCGAGGTGGCMAATGTAGATCCAAAGCCTTGAAGCCACGCTAG 3622
QY 453 ---AsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTyrPheArgHis 471
DB 3623 AGAACTTCTTACCGAGAGAGTGAAGCTGAAGCAGCGTGCCTGACCTGAG----- 3673
QY 472 TyrAspPheArgLysSerCysThrValArgAlaIleSerGlnGluAspGlnLeuAlaPro 491
DB 3674 -----AGGTGAGAGGTGCGCTGCTGACAG----- 3697
QY 492 CysThrGlnAlaGluHisProAlaGluArgValArgSerAlaPheAlaAlaProSerThr 511

```


Db 3698 -----ATGTTGAGAGAGCTGCCGAGGCAAGCGCCCGCCAGCACT 3739
Qy 512 AspSerArgGlnGlyThrProPheArgAlaArgAspGlnPro 526
Db 3740 CCAGAGCCAGACTGCACCCAGCTGGAGCCCAAGCGATTGACCG 3784

RESULT 15
US-09-762-311-3

/ Sequence 3, Application US/09762311
/ Patent No. 6825004
/ GENERAL INFORMATION:
/ APPLICANT: BLUMENFELD, Marta
/ APPLICANT: BOUGUELERET, Lydie
/ APPLICANT: CHUMAKOV, Ilya
/ TITLE OF INVENTION: Nucleic Acids Encoding Human TBC-1 Protein And Polymorphic Markers
/ TITLE OF INVENTION: Thereof.
/ FILE REFERENCE: 46.US2.PCT
/ CURRENT APPLICATION NUMBER: US/09/762,311
/ CURRENT FILING DATE: 2001-06-21
/ PRIOR APPLICATION NUMBER: US 60/095,653
/ PRIOR FILING DATE: 1998-08-07
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: Patent.pm
/ SEQ ID NO 3
/ LENGTH: 3983
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 171..3725
/ NAME/KEY: polyA_signal
/ LOCATION: 3942..3947
/ OTHER INFORMATION: AATAAA
/ NAME/KEY: misc_feature
/ LOCATION: 36
/ OTHER INFORMATION: n=a, g, c or t
US-09-762-311-3

Alignment Scores:
Pred. No.: 0.0258 Length: 3983
Score: 175.50 Matches: 59
Percent Similarity: 42.2% Conservative: 41
Best Local Similarity: 24.9% Mismatches: 114
Query Match: 5.9% Indels: 23
DB: 3 Gaps: 4

US-10-071-838-2 (1-549) x US-09-762-311-3 (1-3983)

Qy 85 TrpGluValTyrLysSerSer-----ArgLysLeu 94
Db 2490 TGGGAAAAGATGCTTAGCACTCCAGAGAGATCAAAAATTAGTTTGACATGGAAAAAATG 2549
Qy 95 IleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeu 114
Db 2550 CACTCGGCTGTGGGCAAGGTGTGCCACGTCAATCACCGAGTGAAATCTGAAATTCTTA 2609
Qy 115 LeuAsnIleGlnGluMetLysLeuLysAsnProGlyArgTyrGln-----Ile 130
Db 2610 GCTGAGCAATTCACCTTAACACACAGTTTCCAGCAACACAGCCAAAGGATGTGCCA 2669
Qy 131 MetLysGluLysGlyLysArgSerSerGluHisIleGlnArgIleAspArgAspValSer 150
Db 2670 TACAAAGAACTTTAAAGCAGCTGATTCGCCAGCAGCAGCGATTTCTATTGACCTTGGG 2729
Qy 151 GlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeu 170
Db 2730 CGAACCTTTCTTACACACCATACTCTCTGCCAGCTTGGAGCAGACAGCTATCGCTT 2789
Qy 171 LeuHisIleLeuLeuAlaTyrGlnGluTyrAsnProGluValGlyTyrCysArgAspLeu 190
Db 2790 TACAACATTTTGAAGGCGCTACTCACTTCTAGACCAAGGAAGTGGAATATTGCCAAGTCTC 2849

Qy 191 SerHisIleAlaAlaLeuPheLeuLeuTyrIleuProGluLysAspAlaPheTrpAlaLeu 210
Db 2850 AGCTTTGTAGCAGGCAATTTGCTTCTTCATATGAGTAGAGAAAGCGCTTAAATGCTC 2909
Qy 211 ValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGly 230
Db 2910 AAGTTCTGATGTTTAC-----ATGGGCTGCGGAAACAGTATCGGCCAGACATGATT 2963
Qy 231 ThrValGlnGlyLeuGlnAspGlnGlnGlnHisValAlaIleThrSerGlnProLysThr 250
Db 2964 ATTTTACAGATCCAGATGTACCAAGCTCTCGAGGTTGCTTCATGATTACCAAGACCTC 3023
Qy 251 MetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIle 270
Db 3024 TACAATCACCTTGAGAGACACGAGATCGGCCCCCAGCTTACGCTGCCCTGCTTCCTC 3083
Qy 271 ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeu 290
Db 3084 ACCATGTTTGCTTCACAGTTCCCGCTGGGATTCTGTAGCCAGAGTCTTTGATATGATTTT 3143
Qy 291 ValGlnGlyGlnAlaLeuMetProIleThrArgIleAlaPheLysVal 307
Db 3144 CTTCAAGGA-----ACAGAGTCAATATTAAAGTG 3173

Search completed: April 5, 2006, 14:54:45
Job time : 305 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:42:42 ; Search time 6211 Seconds
(without alignments)
4135.583 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983
Sequence: 1 MDVVEVAGSWWAQEREDILM.....TSGPCLGLHLESSQPPPGF 549

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 segs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-O=/abs8/ABSSWEB_spool/US10071838/runat_05042006_141915_5173/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808
-USER=US10071838@CGN_1_1_5315@runat_05042006_141915_5173 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hic: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_g881: *
10: gb_g882: *
11: gb_g883: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2963	99.3	2072	4	AL136860 Homo sapi
2	2820	94.5	2497	4	BC047739 Homo sapi
3	1513	50.7	915	5	BQ923700 AGENCOURT
4	1455	48.8	954	5	BU526609 AGENCOURT
5	1440	48.3	909	3	BI753688 AGENCOURT
6	1397.5	46.8	868	6	CD653466 AGENCOURT
7	1342	45.0	881	5	BU542901 AGENCOURT

8	1338.5	44.9	1013	3	BM560320	BM560320 AGENCOURT
9	1335	44.8	919	3	BI522600	BI522600 603175612
10	1324	44.4	930	5	BQ920917	BQ920917 AGENCOURT
11	1323	44.4	1141	3	BM553146	BM553146 AGENCOURT
12	1302.5	43.7	1079	3	BM563571	BM563571 AGENCOURT
13	1272.5	42.7	831	3	BI828272	BI828272 603078163
14	1243	41.7	709	2	BI088323	BI088323 602852793
15	1240.5	41.6	946	5	BQ891586	BQ891586 AGENCOURT
16	1211	40.6	3109	4	BC037244	BC037244 Homo sapi
17	1181.5	39.6	839	6	CA487310	CA487310 AGENCOURT
18	1181	39.6	964	3	BM914524	BM914524 AGENCOURT
19	1171.5	39.3	840	3	BI829764	BI829764 603079835
20	1168	39.2	667	3	BI827136	BI827136 603077337
21	1162.5	39.0	882	3	BI520522	BI520522 603071531
22	1162	39.0	634	7	CN262008	CN262008 170005977
23	1138	38.1	839	3	BI601698	BI601698 603249940
24	1118	37.5	818	3	BM843014	BM843014 603247059
25	1117	37.4	629	3	BM843014	BM843014 K-EST0120
26	1117	37.4	630	3	BM842553	BM842553 K-EST0119
27	1117	37.4	727	1	AL042933	AL042933 DKFZp434K
28	1116.5	37.4	848	5	BQ942166	BQ942166 AGENCOURT
29	1113	37.3	628	3	BM762465	BM762465 K-EST0043
30	1112	37.3	742	3	BI832051	BI832051 603075809
31	1111	37.2	986	5	BX346181	BX346181 BX346181
32	1097	36.8	628	8	DN999748	DN999748 TC104423
33	1067	35.8	635	1	AV650439	AV650439 AV650439
34	1065	35.7	633	8	DR006725	DR006725 TC105000
35	1056	35.4	802	7	CR766757	CR766757 DKFZp468D
36	1049	35.2	650	3	BI830812	BI830812 603080926
37	1026.5	34.4	979	3	BI830468	BI830468 603073339
38	994	33.3	573	3	BM842932	BM842932 K-EST0120
39	950	31.8	539	5	BX281608	BX281608 BX281608
40	932	31.2	545	7	CN262009	CN262009 170006001
41	931.5	31.2	1045	3	BM563007	BM563007 AGENCOURT
42	925	31.0	940	5	BQ709135	BQ709135 AGENCOURT
43	904	30.3	498	5	BX279913	BX279913 BX279913
44	891	29.9	671	3	BI828386	BI828386 603078333
45	887.5	29.8	875	9	AQ788033	AQ788033 HS_3100_A

ALIGNMENTS

RESULT 1
HSM801828
LOCUS HSM801828 2072 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; CDNA DKFZp434P2235 (from clone DKFZp434P2235).
ACCESSION AL136860
VERSION AL136860.1 GI:12053224
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 2072)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the CDNA sequencing consortium of the
German Genome Project. This clone (DKFZp434P2235) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434P2235
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

Location/Qualifiers

```

SOURCE
1. .2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434P2235"
/db_xref="taxon:9606"
/clone="DKFZp434P2235"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="TBC1 domain family, member 3"
1. .2072
/gene="DKFZp434P2235"
99. .1748
/gene="DKFZp434P2235"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66794.1"
/db_xref="GI:12053225"
/db_xref="UniProt/Swiss-Prot:Q8IZP1"
/translation="MDVVEVAGSWAQEREDIMKYEKHRAGLPEDKPKPFRSYNN
NVDHLGIHETELPPLTAREAKQIRERISKSKWMDMLGWEKYKSRKLIIDRAVKG
PMNIRGPMWSVLNTEEMKLNKPGROYIMKEKSKSEHIQRIIDRVSGTLRKHIFER
DRYGTQKRELHILLAYEENPEVGYCRDLSHIALFLYLPEEDAFWALVQLASER
HSLQGFHSFNGTVQGLQDOEHVAVTSQPTMGHODKDLGQCSPLGCLIRIIDG
ISLGLTLRLMDVYLVEGEOLMPTIRIAFKVQOKRLTKTSRGPWFRCNRFVDTWAR
DEDTVLKLRLASMKKLTTRKGDLPRAKPEQSSASRPVPSRGGKTLCKGDRAAPG
PPAPRPPIWSASPRAPRSTPCPGAVREDTYPVGTGVPSPALAOGGPOGSWRPL
QWNSMPRLPTDLIDVEGPFWRHYDFRQSCWVRAISQEDQLAPCWQAEHPABRVRSAPAA
PSTDSDQGTFRARDEQQCAPTSGLCLHLESSQFPGP"

ORIGIN
Alignment Scores:
Pred. No.:      1.28e-194      Length:      2072
Score:          2963.00        Matches:      545
Percent Similarity: 99.6%      Conservative: 2
Best Local Similarity: 99.3%      Mismatches:  2
Query Match:    99.3%      Indels:        0
DB:              4          Gaps:          0

US-10-071-838-2 (1-549) x HSM801828 (1-2072)
```

```

QY      1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db      99 ATGGAAGTGTAGAGGTCGGCGGCAAGTTGGTGGGCACAAAGACGAGGACATCATATG 158
QY      21 LysTyrGluLysGlnHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db      159 AATAACGAAAAGGACACGAGCTGGGCTGCCAGGACAAAGGGGCTTAAGCCTTTTCCA 218
QY      41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db      219 AGCTACAAACAACAACGTCGATCTTTGGGATGTGATGACGACGAGCTGCCTCTCTG 278
QY      61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db      279 ACTGCGCGGGAAGCGAAGCAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGGTGAT 338
QY      81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db      339 ATGCTGGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCATAGATCGACGTACAAG 398
QY      101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
Db      399 GGAATGCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAACTAGAGGAATG 458
QY      121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db      459 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGTCAATCTGAG 518
QY      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db      519 CACATCCAGCGCATGACCGGGAGCGTAAGCGGACATTAAGGAAGCATATATTCTTCAGG 578
```

```

QY      161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db      579 GATCGATACCGAACCAGACGCGGAATACTCCACATCTCTGGCATATGAGAGTAC 638
QY      181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
Db      639 AACCCGAGGTGGCTACTGTGAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 698
QY      201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db      699 CTTCCTGAGGAGATGATCTTGGGCACTGGTGACAGTGTGGCCAGTGAAGGCACTCC 758
QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db      759 CTGCAGGATTTCAACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 818
QY      241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db      819 CATGTGTAGCCACGTCACACACCAAGACCATGGGGCATCAGACAAAGAAATCTATGT 878
QY      261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db      879 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTGATGACGGGATCTCTCCGG 938
QY      281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db      939 CTCACCTGCGCTGTGGACGTGTATCTGTGTAAGGCGAAGCAGCGCTGATGCCGATA 998
QY      301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db      999 ACAAGATCGCCTTAAAGTTACGACAAAGCGCTCACGAAGACGTCAGGTGTGGCCG 1058
QY      321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db      1059 TGGGACAGTTTTCGAACCGGTTCGTTGATACCTGGCGCAGGATGAGACACTGTCTC 1118
QY      341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db      1119 AAGCATCTTAGGCTCTTATGAAGAACTAAACAAGAAAGGGGACCTGCCACCCCA 1178
QY      361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db      1179 GCCAAACCCGACAGGGTCTCGGCATCCAGGCCCTGTGCGGCTTCACGTGGCGGGAAG 1238
QY      381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db      1239 ACCCTCTGAAAGGGGACAGCAGGCCCTCCAGGCCACCAAGCCGGTTCCCGGCCC 1298
QY      401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db      1299 ATTTGTCAAGTTCCTCCGCCACGGCACTGTTCTTCACACCTGTCTGTGGGCT 1358
QY      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db      1359 GTCCGGGAAGACACTAACCTGTGGGCACTCAGGCTGCCAGCCCGGCTGGTCTCAG 1418
QY      441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db      1419 GAGGACCTCAGGGTCTCTGGAATCTCTGCAAGTGAACCTCCATGCCCGCTCCCAAG 1478
QY      461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db      1479 GACCTGACGTAGAGGGCCCTTGTGTCCGCAATTATGATTTCAGACAGAGCTGTGGTTC 1538
QY      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db      1539 CGTGCCATATCCAGAGAGACCAAGCTGGCCCTGTGCAAGCTGAACACCCCTGGCAG 1598
QY      501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db      1599 CGGGTGAGATCGGCTTTCGCTGCAACCGACACTGATTCGACAGAGGGGACCCCTTCAAG 1658
```

```

Oy      521  ALaargAspGlUGInProCySaLaProThrSerGlyProCySLeuCySGLyleuHisLeu 540
          |||
Db      1659  GCTAGGAGCAGAACAGCAGTGTCTCCACCTCAGGGCCTTGCTGCGGCTCCACTTG 1718
          |||

Oy      541  GluSerSerGlnPheProGlyPhe 549
          |||
Db      1719  GAAAGTCTCAGTTCCTCCAGGCTTC 1745
          |||

RESULT 2
BC047739      2497 bp      mRNA      linear      HTC 20-MAR-2003
LOCUS      BC047739
DEFINITION      Homo sapiens, similar to Rab GTPase-activating protein PRG17, clone
IMAGE:5743752, mRNA.
ACCESSION      BC047739
VERSION      BC047739.1  GI:29126829
KEYWORDS      HTC.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
1 (bases 1 to 2497)
Strausberg,R.
Direct Submission
Submitted (03-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REFERENCE      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
Email: cgagbs.rem@nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site:      http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxll.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 98 Row: d Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984
This clone has the following problem: retained intron.
FEATURES
source
1.2497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5743752"
/tissue_type="Brain, adult medulla"
/clone_id="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: PCMV-SPORTc"

ORIGIN

Alignment Scores:
Pred. No.:      1.32e-184      Length:      2497
Score:      2820.00      Matches:      542
Percent Similarity:      79.0%      Conservative:      2
Best Local Similarity:      78.7%      Mismatches:      5
Query Match:      94.5%      Indels:      140
DB:      4      Gaps:      2

US-10-071-838-2 (1-549) x BC047739 (1-2497)

Oy      1  MetAspValValGluValAlaGlySerTrpAlaGlnGluArgGluAspIleIleMet 20
          |||
Db      89  ATGACGCTGGTAGAGGTGCGCGGCGAGTGTGGGCAAGAGCGAGAGCATCATATG 148
          |||

```

QY	21	lySTYrGlulYsGlyhIaArgAlaGlyLeuProGluAspIlySGlyProIysProPheArg	40
Db	149	AAATACGAAAGGACACCCGAGCTGGCTGCCAGAGCAACAGGGGCTTAAGCTTTTCGA	208
QY	41	SerTYrAsnAsnValAspHisIleuGlyIleVal-----	52
Db	209	AGCTACAACAACAACGTCGATCATTTGGGGATTGTACAGTCGGGTCAACCTCTTCCAG	268
QY	52	-----	52
Db	269	GGACAGAACTCTCCCTGACTCCCTGCAGGTCACGCCGAGTTGTTAGCCAGAGTG	328
QY	52	-----	52
Db	329	TGGGGCCCATTAAGGAGCCGGTGGAAATGGAAGACTGGGCTAGGTCAAGCCCCCTGGCGC	388
QY	52	-----	52
Db	389	TCAGCAGTTCTGTCCGCAAGTAGACACAAGAGCGGGGACGCTGAGGGTCTGGCCCT	448
QY	53	-----HisGluThrGluLeuProProLeuThrAlaArgGluAlaIy	66
Db	449	GTCTACTTGAGACAACCCCGTGAGACGAGCTGCCTCTCTGACTGCGCGGAGCGCAA	508
QY	66	sgInIleaArgGluIleSerArgIysSerIySTrValAspMetIeuGlyAspTrpG	86
Db	509	GCAAAATTCGGCGGAGATCAGCCGAAAGACAAGTGAGTGATGTCTGGAGACTGGGA	568
QY	86	ulYsTYrIysSerSerArgIySleuIleAspArgAlaTYrIySGlyMetProMetAsnI	106
Db	569	GAAATACAACAAGCAGACGAAAGCTCATAGATCGAGCGTACAAGGGAATGCCCATGAACAT	628
QY	106	eArgGlyProMetTrpSerValIleuIleuAsnIleGluGluMetIySleuIyAsnProG	126
Db	629	CCGGGGCCGATGTGGTCAGTCTCTGAAcATTGAGAAATGAAGTTGAAAAACCCCG	688
QY	126	yArgTYrGlnIleMetIySGlyIySlyArgSerSerGluHisIleGlnArgIleAs	146
Db	689	AAGATACCAAGATCATGAAGAGAGGCAAGAGTCATCTGACACATCCAGCGCATGCA	748
QY	146	pArgAspValSerGlyThrIleuArgIyHisIlePhePheArgAspArgTYrGlyThrIy	166
Db	749	CCGGGACGTAAGCGGACATTAAGGAACATATATTCTTCAAGGATCGATACGGAACCA	808
QY	166	sgInArgGluIleuIleHisIleIleuIleuAlaTYrGluGluTYrAsnProGluValGlyTy	186
Db	809	GCAGCGGAATACTCCACATCTCTCGCATATGAGAGTAACAACCGAGGTGGCTA	868
QY	186	rCyAspAspIleuSerHisIleAlaIleuPheIleuIleuTYrIleuProGluIleuAspAl	206
Db	869	CTGACAGGACCTGAGCCACATCGCGCCTTGTCTCTCTATCTTCCCTGAGAGAGATGC	928
QY	206	aPheTrpAlaIleuValGlnIleuIleuAlaSerGluArgHisSerIeuGlnIyPheHisSe	226
Db	929	ATTCTGGGCACTGTGTGACGCTCTGGCCAGTGAGGACCTCCCTGCAGGATTTCACAG	988
QY	226	rProAsnGlyGlyThrValGlnIyIleuGlnAspGlnGlnIleHisValAlaIleThrSe	246
Db	989	CCCAATGCGGGACCGCTCCAGGGGCTCCAAGACCAACAGAGCATGTGTAGCCACGTC	1048
QY	246	rGlnProIySThrMetGlyHisGlnAspIySlyAspIleuCySGlyIleCySerProIe	266
Db	1049	ACAACCCCAAGACCATGGGCACTCAGGACAAGAAAGATGTATGTGGGCAGTGTTCCCGTT	1108
QY	266	uGlyCysIleuIleArgIleIleuIleAspGlyIleSerIeuGlyIleuThrIleuArgIleuTr	286
Db	1109	AGGCTGCTCATCCGATATTGATTGACGGGATCTCTCTCGGGCTCACCCCTGGCTGTG	1168
QY	286	pAspValTYrIleuValGluGlyGluGlnAlaIleuMetProIleThrArgIleAlaPheIy	306
Db	1169	GGAACGTATCTGTAGAAAGCGAACAGGCGCTGATGCCGATTAAcAAAGAAATCGCTTTAA	1228
QY	306	sValGlnGln-----	309

Db	1229	GGTTCAGCAGAGTAAGTCTACGTGTGCCACGGGGCCCTGGGGAGCCCTGGGGTCAAGACC	1288
QY	309	-----	309
Db	1289	CCGACTGGCCCGAGGGCAGCTTCCTCACACTGTCTCATGATCTCTGTCTTGCGCCAGAGA	1348
QY	309	-----	309
Db	1349	GGGAGGTCTGGCCAGGTGGGCTGGGAGAGACATGTGACACCGAGCCCATCCCCACATG	1408
QY	310	-----LyAsrgLeuThrLysThrSerArgCysGlyProTrpA :::	322
Db	1409	ACCCAGATGAAGAATCGAGAGTGTGAGCGCCTCAGGAAGACGTCCAGGTGTGGCCGTGGG	1468
QY	322	laArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLysH	342
Db	1469	CACGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGATGAGGACACTGTGCTCAAGC	1528
QY	342	IsLeuArgAlaSerMetCysLysLeuThrArgLysGlnGlyAspLeuProProProAla	362
Db	1529	ATCTTAGGGCCTTATGAGAAACTAACCAAGAAAGGGGAGCTTGCCACCCCAAGCCA	1588
QY	362	YsProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLysThrL	382
Db	1589	AACCCGAGCAAGGTCGTCGGCATCCAGGCGCTTGCCGCTTCACGTGGCGGGAAGACCC	1648
QY	382	euCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgProIleT	402
Db	1649	TCTGCAAGGGGAGCAGGAGGCCCTCCAGGCCCAACGCCGCTTCCCGCGGCCCATTT	1708
QY	402	rpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAlaValA	422
Db	1709	GGTCAAGCTTCCCTGCCACCGGGACCTCGTCTTCACACCCGTGCTGGTGGGGCTGTCC	1768
QY	422	rgGluAspThrTyrrProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGly	442
Db	1769	GCGAAGACACCTACCTGTGGGCACTCAGGGGTGGCCCAAGCCCGGCTTGCTCAGGGAG	1828
QY	442	lyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThrAspL	462
Db	1829	GACCTCAGGGTTCCTGAGATTCTGTGAGTGAATCCATGCCCCGCTTCCCAACGGACC	1888
QY	462	euaAspValGlnGlyProTrpPheArgHisTyrrAspPheArgGlnSerCysTrpValArgA	482
Db	1889	TGGACGTAGAGGGCCCTGGTTCGCCCATTTATGATTTCAGACAGAGCTGTGGGCCGTG	1948
QY	482	laIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGluArgV	502
Db	1949	CCATATAACAGAGGAGCACGCTGGCCCCCTGTGGCAGGCTGAACACCCCTGGGAGCGGG	2008
QY	502	alArgSerAlaPheAlaIleAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaA	522
Db	2009	TGAGATCGGCTTTCGCTGCACCCACGACTGATTCCGACCAAGGGCACCCCTTCAGAGCTA	2068
QY	522	rgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeuGlnS	542
Db	2069	GGGACGAACAGCAGTGTGTCTCCACCTCAGGGGCTTGCTTGCGGCTTCACATTGGAAA	2128
QY	542	exSerGlnPheProProGlyPhe	549
Db	2129	GTTCTCAGTTCCTCCAGGCTTC	2151
RESULT 3			
LOCUS	BQ923700	915 bp	mRNA
DEFINITION	AGENCOURT_8798484 NIH_MGC_101 Homo sapiens CDNA clone IMAGE:6425622		EST 20-AUG-2002
ACCESSION	BQ923700		
VERSION	BQ923700.1	GI:22338731	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 915)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
plate: LLCM2607 row: 1 column: 07
High quality sequence stop: 639.

FEATURES
Source
1. 915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6425622"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_11b="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Alignment Scores:
Pred. No.: 1.13e-94 Length: 915
Score: 1513.00 Matches: 295
Percent Similarity: 94.6% Conservative: 2
Best Local Similarity: 93.9% Mismatches: 6
Query Match: 50.7% Indels: 12
DB: 5 Gaps: 1

US-10-071-838-2 (1-549) x BQ923700 (1-915)

QY	137	ArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHis	156
Db	1	AAGCATCTGAGCAGATCCAGCGCATGACCGGAGCGTAAGCGGGACATTAAAGAAAGCAT	60
QY	157	IlePhePheArgAspArgTyrrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAla	176
Db	61	ATATTCTTCAGGAGTGAATACGGAACCAAGCAGCGGAATCTACTCCATCTCTGGCA	120
QY	177	TyrrGlnGluTyrrAsnProGluValGlyTyrrCysArgAspLeuSerHisIleAlaLeu	196
Db	121	TATGAGGAGTATAACCGGAGGTGGCTACTGCAAGGACCTGAGCCATCGCGCTTG	180
QY	197	PheLeuLeuTyrrLeuProGlnGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSer	216
Db	181	TTCCTCTCTATCTTCTTGAGGAGGATGCAATTCTGGGCACCTGTGACGTGCTGGCAGT	240
QY	217	GluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGln	236
Db	241	GAGAGGCACTCCCTGAGGGGATTTCAACAGCCCAATGGCGGACCTCCAGGGGCTCCAA	300
QY	237	AspGlnGlnGluHisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLys	256
Db	301	GACCAACAGAGCATGTGTAGCCACGTCACCAACCAAGACCATGGGCATCAGACAAG	360
QY	257	LyAspLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGly	276

Db 361 AAAGATCTATGTGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTGATTGACGGG 420

Qy 277 ILeSerLeuGlyLeuThrLeuArgLeuTyrPaspValTyrLeuValGluGlyGluAla 296

Db 421 ATCTCTCTCGGGCTCACCTCGCGCTGTGGACGTATCTGTAGAACGCGAACAGCGG 480

Qy 297 LeuMetProIleThrArgIleAlaPheIysValGlnGlnIlybArgLeuThrIlybThrSer 316

Db 481 TTGATGCCGATACAGAAATCGCCTTTAAGTTCAAGACAGCGCCTTCAAGAACCTCC 540

Qy 317 ArgCybSgLyProTyrPalaArgPheCybAsnArgPheValAspThrTyrPalaArgAspGlu 336

Db 541 AGGTGTGGCCCGTGGGCACTTTTGCAACCGGTTCTTGATACCTGGGCCAGGATGAG 600

Qy 337 AspThrValLeuIlybHisLeuArgAlaSerMetIlybLybLeuThrArgLybGlnIlybAsp 356

Db 601 GACACTGTGCTCAAGCATCTTANGCCCTGTAGAAAGAACTAACAGAAAGACAGGGGAC 660

Qy 357 LeuProProProAlaIlybProGluGlnIlybSerSerAlaSerArgProValProAlaSer 376

Db 661 CTGCAACCCCAAGCAAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCA 720

Qy 377 ArgGlyGlyIlybThrLeuCybIlybSgLyAspArgGlnAlaProProGlyProProAlaArg 396

Db 721 CGTGGCGGGAAGACCTCTGCAAGGGGAGACAGGCCCTCCAGC-CCACCAAGCCCGG 779

Qy 397 PheProArgProIleTyrPserAlaSerProProAlaArgAlaProArgSerSerThrProCyb 416

Db 780 TTCCCGCGGCCCATTTGGTTCAGCTTCCCGCCACGGGACCTCGTTCTTCCACACCTGT 839

Qy 417 -ProGlyGlyAlaValArg-GluAspThrTyr-ProValGlyThrGlnIlybValProSer 435

Db 840 TCCTGTGGGGCTGTCCGGGGAAGACACTAACCTCGGGGGCACTCAAGGTGTGCCCCAN 899

Qy 436 ProAlaLeuAlaGlnGlyGlyProGlnIlybSerTyr 447

Db 900 CCC-----GGGCCCTGG 911

RESULT 4

LOCUS BUS26609 954 bp mRNA linear EST 13-SEP-2002

DEFINITION AGENCOURT 10181753 NIH-MGC 101 Homo sapiens cDNA clone

IMAGE:6536307 5', mRNA sequence.

ACCESSION BUS26609

VERSION BUS26609

KEYWORDS BUS26609.1 GI:22837050 EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 954)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLCM2697 row: h column: 03

High quality sequence stop: 627.

FEATURES

Source

1. 954

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6536307"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_1ib="NIH-MGC 101"

/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 1.24e-90 Length: 954

Score: 1455.00 Matches: 283

Percent Similarity: 94.4% Conservative: 2

Best Local Similarity: 93.7% Mismatches: 11

Query Match: 48.8% Indels: 6

DB: 5 Gaps: 2

US-10-071-838-2 (1-549) x BUS26609 (1-954)

Qy 173 ILeLeuLeuAlaTyrGluGluTyrAsnProGluValGlyTyrCybArgAspLeuSerHis 192

Db 1 ATCTCTCTGGCATATGAGAGATATAACCCGAGGTGGGCTACTGACGGACCTGAGCCAC 60

Qy 193 ILeAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPheTyrPalaLeuValGln 212

Db 61 ATGCGCGCTGTCTCTCTCTCTATCTTCTGAGAGATGATCTGGGCACTGTGTGACG 120

Qy 213 LeuLeuAlaSerGluArgHisSerLeuGlnIlybPheHisSerProAsnGlyGlyTyrThrVal 232

Db 121 CTGCTGGCCAGTGAAGGCACTCCCTGAGGAGATTTCACAGCCCAATGGCGGACCGTC 180

Qy 233 GlnGlyLeuGlnAspGlnGlnIlybHisValAlaThrSerGlnProIlybThrMetGly 252

Db 181 CAGGGGCTCCAGAACCAACAGAGCATGTGTAGCCACGTCAACCAAGACCATGGGG 240

Qy 253 HisGlnAspIlybLybAspLeuCybGlyGlnCybSerProLeuGlyCybLeuIleArgIle 272

Db 241 CATCAGGACAGAAAGATCTATGTGGGAGTGTCCCGCTTACGCTCATCCGATTA 300

Qy 273 LeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTyrPaspValTyrLeuValGlu 292

Db 301 TTGATTGACGGGATCTCTCTCGGGCTTCACTCGCGCTGTGGGACGTGTATCTGTAGA 360

Qy 293 GlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheIysValGlnIlybArgLeu 312

Db 361 GGCGAACAAGCGTTGATGCCGATACAGAATCGCCTTAAGTTCAAGCAAGCGCCTC 420

Qy 313 ThrIlybThrSerArgCybGlyProTyrPalaArgPheCybAsnArgPheValAspThrTyr 332

Db 421 ACGAAGACGTCCAGGTGTGGCCCGTGGCAGCTTTTGCAACCGGTTCTGTGATACCTGG 480

Qy 333 AlaArgAspGluAspThrValLeuIlybHisLeuArgAlaSerMetIlybLybLeuThrArg 352

Db 481 GCCAGGGATGAGACACTGTGTCAAGCATCTTAGGGCCTTATGAAGAACTAACACAGA 540

Qy 353 IlybGlnIlybAspLeuProProProAlaIlybProGluGlnIlybSerSerAlaSerArgPro 372

Db 541 AAGCAGGGGGAAGCTGCAACCCCAAGCAAAACCGAGCAAGGTGTGTGGCATCCAGGCT 600

Qy 373 ValProAlaSerArgGlyGlyIlybThrLeuCybIlybSgLyAspArgGlnAlaProProGly 392

Db 601 GTGCCGGCTTCACTGTGGCGGAAGACCTCTGCAAGGGGAGACAGGACCCCTCAGGC 660

Qy 393 ProProAlaArgPheProArgProIleTyrPserAlaSerProProArgAlaProArgSer 412

Db 661 CCACCAAGCCCGGGTCCCGCGCCCAATTGTGTCAGCTTCCCGCCACGGCACCTGTTCT 720

Qy 413 SerThrProCybProGlyGlyAlaValArgGluAspThrTyrProValGlyThrGlnIly 432

Db 721 TTCACACCTGTCTGTGTGGGGCTGTCCGGGAAACACTTAACCTGTGTGGCACTCAGGGT 780

```

Qy      433 ValProSerProAla-LeuAlaGlnGlyGly-ProGlnGlySerTyrParPhenLeu--G1 451
      |||
Db      781 GTGCCAGCCCGCCCTGCTCAAGAGGAACCTCAGGGTTCTGGAGATCCTGGCAG 840
Qy      451 nTTPAsnSerMet--ProArGleuProThrAspLeuAspVal---GluGlyProTyr 468
      : |||
Db      841 GGAAACTCCATGGCCCCCGCTCCCAACGAGCCTGGACGTTAAAGGGGGCCCTGG 898

RESULT 5
BI753688
LOCUS      909 bp      mRNA      linear      EST 25-SEP-2001
DEFINITION 603023590F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193959 5',
            mRNA sequence.
ACCESSION  BI753688
VERSION    BI753688.1  GI:15745266
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 909)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strauberg, Ph.D.
            Email: cga@bs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.lnl.gov
            plate: LLM11485 row: d column: 24
            High quality sequence scop: 884.

FEATURES
     source
         1..909
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:5193959"
             /lab_host="DH10B"
             /clone_lib="NIH_MGC_114"
             /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
             Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
             male brains, age range 23-27 yo. Library is oligo-dT
             primed and directionally cloned (EcoRV site is destroyed
             upon cloning). Average insert size 1.5 kb, insert size
             range 1-3 kb. Library is normalized and enriched for
             full-length clones and was constructed by C. Gruber
             (Invitrogen). Research Genetics tracking code 019. Note
             this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      1.27e-89      Length:      909
Score:          1440.00      Matches:      288
Percent Similarity: 94.8%      Conservative: 3
Best Local Similarity: 93.8%      Mismatches: 9
Query Match:    48.3%      Indels:      7
DB:             3          Gaps:      0

US-10-071-838-2 (1-549) x BI753688 (1-909)

Qy      56 GluLeuProProLeuThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLys 75
      |||
Db      3  GAGTCGCTCTCTGACTGCGCGGAGCGCAAGCAAAATTCGGCGGAGATCAGCCGAAGA 62
Qy      76 -SerLysTyrValAspMetLeuGlyAspTyrGluLysTyrLysSerSerArgLysLeuI1 95
      |||
Db      63 GAGCAAGTGGGTGATATGCTGGAGACTGGAGAAATACAAAGCAGAGAAAGCTCAT 122

```

Qy	95	eAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLe	115			
Db	123	AGATCGAGCGTACAGGGGAATGCCATGAACATCCGGGGCCCGATGTGTCAGTCTCCT	182			
Qy	115	uAsnIleGluGluMetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysG	135			
Db	183	GAACACGTAGGAATGAAGTTGAAAAACCCCGAAGATACCAGATCATGAAGAGAAAGG	242			
Qy	135	YLysArgSerSerGluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLy	155			
Db	243	CAAGAGGTCACTGAGCACATCCAGCGCATCGACCGGGACGTAAAGCGGCATTAAGAA	302			
Qy	155	SHisIlePhePheAlaArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeu	175			
Db	303	GCATATATTCTTCAGGGATCGATACGGAAACCAAGCAGCGGGAACCTACTCCACATCTCCT	362			
Qy	175	uAlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaI	195			
Db	363	GGCATATGAGAGTATAACCCGAGGTGGGCTACTGACGGACCTGAGCCACATCGCCGC	422			
Qy	195	AlaPheLeuLeuTyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAl	215			
Db	423	CTTGTTCCCTCTATCTCTCTGAGAGAGATGCATCTGGGCACGTGGTCAGCTGCTGGC	482			
Qy	215	ASerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLe	235			
Db	483	CAGTAGAGGCACTCCCTGCAGGGAATTTCACAGCCCAATGGCGGACCGTCCAGGGGCT	542			
Qy	235	uGlnAspGlnGlnGluHisValAlaIleHisSerGlnProLysThrMetGlyHisGlnAs	255			
Db	543	CCAAGACCAACAGAGCATGTGTAGCCACTCAACAACCAAGACCATGGGSCATCAGGA	602			
Qy	255	PLysLysAspLeuCybGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAs	275			
Db	603	CAAGAAAGATCTATGTGGGCAGTGTCCCGCTTAAGCTGCCTCATCCGATATCGATTGA	662			
Qy	275	pGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGluI	295			
Db	663	CGGATCTCTCTCGGGCTCACCCCTCGCGCTGTGGGACGTATCTGTAGAAAGCGAACA	722			
Qy	295	nAlaLeuMetPro-IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysT	315			
Db	723	GGCGTAGATGCCGAATTACAAGAATCGCCTTTAAGTTCAAGCAGAAAGCGCTCACGAAGA	782			
Qy	315	hr-SerArgCysGlyProTrpAla--ArgPheCysAsnArgPheValAspThrTrp-Ala	333			
Db	783	CGTTCACGGTGTGGCCCCGTGGGGCCACGTTTGTGCAACGGGTTCTGTAGATACTGGGGCC	842			
Qy	334	ArgAsp-GluAspThrValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLy	353			
Db	843	AGGCATTGAGGACACTGTGCTCAAGCATCTTAGGGCCTCTATGAAGAACTAACAGAAA	902			
Qy	353	eGlnGly 355				
Db	903	GCACGGG 909				
RESULT 6						
LOCUS	CD653466	868 bp	mRNA linear EST 18-JUN-2003			
DEFINITION	AGENCOURT_14539069 NIA Human H1 Embryonic Stem Cell cDNA Library					
ACCESSION	(Long) Homo sapiens cDNA clone IMAGE:30423486 5', mRNA sequence.					
VERSION	CD653466					
KEYWORDS	CD653466.1 GI:31891804					
SOURCE	EST.					
ORGANISM	Homo sapiens (human)					
Homo sapiens						
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;						
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;						
Hominidae; Homo.						
1 (bases 1 to 868)						
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/ .					
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)					
TITLE						

JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@b3-r@mail.nih.gov
Tissue Procurement: Irene Gjinis and Mahendra Rao, NIA
CDNA Library Preparation: Yulan Piao and Minoru Ko
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC C lone distribution information
can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM504 row: j column: 07
High quality sequence stop: 695.
Location/Qualifiers

FEATURES

source

1. 868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30423486"
/issue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA library (Long)"
/note="Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
cell line. Undifferentiated human ES cell line WA01/H1
was obtained from Wicell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed round colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with TRIzol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:11544199]) Double-stranded cDNAs were
synthesized with an Oligo(dt) primer [Invitrogen:
5'-pGACTAGTCTAGATCCGACGCGCCGCTTTTCTTTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lona-linker L1-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
pCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."

ORIGIN

Alignment Scores:
Pred. No.: 1.05e-86 Length: 868
Score: 1397.50 Matches: 267
Percent Similarity: 94.1% Conservative: 2
Best Local Similarity: 93.4% Mismatches: 15
Query Match: 46.8% Indels: 2
DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x CD653466 (1-868)

Qy 151 GlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGluLeu 170
Db 4 GGTAATCCAAAGCAATATATTCTTCAGGATCGATACGGAACCAAGCAGCGGAACCTA 63

Qy 171 LeuHisIleLeuLeuAlaTyrGluGluTyrAsnProGluValGlyTyrCysArgAspLeu 190
Db 64 CTCACATCTCTCGCATATGAGAGTATAACCCGAGGTGGCTACTGCAGGAGACTG 123
Qy 191 SerHisIleAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPheTyrAlaLeu 210
Db 124 AGCCACATCGCCCTTGTCTCTCTTAATCTTCTGAGGAGGATTCATTCGGGACTG 183
Qy 211 ValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyGly 230
Db 184 GTGACGCTGTGGCCAGTGAGAGGCACTCCCTGACGGGATTCACAGCCCAATGGCGG 243
Qy 231 ThrValGlnGlyLeuGlnAspGlnGlnIuHisValValAlaThrSerGlnProLysThr 250
Db 244 ACCGTCCAGGGCTCCAGAACCAAGACATGTGTAGCCACGTCAACACCAAGACC 303
Qy 251 MetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeuGlyCysIleuIle 270
Db 304 ATGGGCGATCAGGACAAGAAAGATCTATGTGGCAGTGTCCCGTTAGGCTGCCTCATC 363
Qy 271 ArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTyrAspValTyrLeu 290
Db 364 CGGATATTGATTGACCGGATCTCTTCGGGCTCACCTCGCCTGTGGACGTGTATCTG 423
Qy 291 ValGluGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLysValGlnGlnLys 310
Db 424 GTAGAGGCGGAACAGCGCTTGATGCGATAACAGAAATCGCTTTAAGGTTACAGCAGAG 483
Qy 311 ArgLeuThrLysThrSerArgCysGlyProTyrAlaArgPheCysAsnArgPheValAsp 330
Db 484 CGCCTACGGAAGACGTCCAGGTGTGGCCCGTGGCAGCAGTTTGTGCAACCGGTCGTTGAT 543
Qy 331 ThrTyrAlaArgAspGluAspThrValLeuLysHisIleuArgAlaSerMetLysIleu 350
Db 544 ACCTGGGCGGAGATGAGACACTGTGCTCAAGCATCTTAGGGCTCTATGAAGAACCTA 603
Qy 351 ThrArgLysGlnGlyAspLeuProProProAlaLysProGluGlnGlySerSerAlaSer 370
Db 604 ACAAGAAAGCAGGGGAGCCTGCACACCCCAAGCCAAACCCAGCAAGGTCGTGGCATCC 663
Qy 371 ArgProValProAlaSerArgGlyLysThrLeuCysLysGlyAspArgGlnAlaPro 390
Db 664 AGGCTGTGTGGCTTCACTGAGGTGGGAGAACCTCTGCAAGGGGAGACAGGCGGCTT 723
Qy 391 ProGlyProProAlaArgPheProArgProIleTyrPheAlaSer-ProProArgAlaPr 410
Db 724 CCAGNCCACCAAGCCGCTTCCGNCAGCAATGTGTAAGTCTCCCGCCACGGGACAC 783
Qy 410 OArgSerSerThrProCysProGlyGlyAlaValAlaArgGluAspThrTyrProValGlyTh 430
Db 784 TCGTCTTCCACACCTGTCTGTGTGGGCTGTG--NGNNAACCTTACCTGTGGGCTT 840
Qy 430 rGlnGlyValProSer 435
Db 841 CAGTGTGCCACCCGCC 856
RESULT 7
BUS42901 881 bp mRNA linear EST 13-SEP-2002
LOCUS BUS42901
DEFINITION AGENCOURT_10334768 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574864
5', mRNA sequence.
ACCESSION BUS42901
VERSION BUS42901.1 GI:22853384
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: L1CM2770 row: n column: 16
High quality sequence stop: 642.

FEATURES

source

1..881
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6574864"
/issue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAAGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 7.48e-83 Length: 881
Score: 1342.00 Matches: 267
Percent Similarity: 91.4% Conservative: 8
Best Local Similarity: 88.7% Mismatches: 17
Query Match: 45.0% Indels: 11
DB: 5 Gaps: 2

US-10-071-838-2 (1-549) x BU542901 (1-881)

QY 81 MetLeuGlyAspTrpGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
:::|||||
DB 1 TTGCTGGGAGACTGGAGAA-TACAAAGCAGCAGAAAGCTCATAGATCAGCGTACAAAG 59
QY 101 GlyMetProMetAsnIleArgGlyProMetTyrSerValLeuValAsnIleGluMet 120
60 GGAAATGCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAAACATGAGGAATG 119
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysGlyLysArgSerSerGlu 140
|||||
DB 120 AAGTTGAAAAAACC CGAAGATACCAAGATCATGAAGAGAAAGGCAAGGTCATCTGAG 179
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
|||||
DB 180 CACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGGAAGCATATATTCTTCAGG 239
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
|||||
DB 240 GATCGATACGGAACCAAGCAGCGGGAATACTCCACATCTCTCTGGCATATGAGAGTAT 299
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr 200
|||||
DB 300 AACCCGAGGTGGGTACTGACAGGACCTGAGCCACATCGCGCTGTCTCTCTAT 359
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
|||||
DB 360 CTTCCTGAGGAGGATGATTCGTGGCACTGTGACAGTCTGGCCAGTGAAGGCACTCC 419
QY 221 LeuGlnGlyPheHisSerProAlaGlnGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
|||||
DB 420 CTGCAGGAGATTTCACAGCCCAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 479
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260

DB 480 CATGTGATGCCACGTACACACCCAGACCATGCGGCATCAGAGACAAGAAAGATCTATGT 539
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 540 GGGCAGTGTCCCGGTAGGCTGCTCATCCGATATTGATGACGGGATCTCTCCGG 599
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
DB 600 CTCACCCCTGCGCTGTGGACGTGTATCTGTAGAAGGCAACAGCGCTGATGCCGATA 659
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 660 ACAAGATCGCCTTTAAAGTTACAGCAAGCGCCTCACGAAAGACGTCCAGGTGTGCCG 719
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrp-AlaArgAspGluAspThrValIle 340
DB 720 TGGGCACGTTTTTGCAACCGGTTCGTTGATACCTGGGCGCAGGATGAAGACACTGTGC 779
QY 340 uLysHis---LeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPr 359
DB 780 TCAAGCAATCTTAAAGGCGCTCTATGAAGAAACCTAACCAAGAAAGCCAGGGGGACC 839
QY 359 oProAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyG1 379
DB 840 CTG-GCAAAACCCACG-----CCACAACCCCGAAACCAAGGGG 877
QY 379 Y 379
DB 878 T 878

RESULT 8
BM560320 1013 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6564065 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744726
DEFINITION 5', mRNA Sequence.
ACCESSION BM560320
VERSION BM560320.1 GI:18804652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1013)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM12766 row: 1 column: 15
High quality sequence start: 36
High quality sequence stop: 745.
FEATURES
source
1..1013
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744726"
/issue_type="medulla"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-Sports; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores: 1.58e-82 Length: 1013
Pred. No.: 1338.50 Matches: 257
Score: 1338.50
Percent Similarity: 93.4% Conservative: 14
Best Local Similarity: 88.6% Mismatches: 15
Query Match: 44.9% Indels: 4
DB: 3 Gaps: 1

US-10-071-838-2 (1-549) x BMS60320 (1-1013)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 98 ATGACGCTGTAGAGGTCGTGGGTAGTTGGTGGGCAACAAGCGAGAGCATCATATTATG 157
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 158 AATATCGAAAGGGACACCGAGCTGGGCTGCCAGAGACAAGGGGCTTAAGCTTTTGA 217
QY 41 SerTyrAsnAsnValAspHisIleuGlyIleValHisGluThrGluLeuProProLeu 60
DB 218 AGCTACAACAACAATCATCTTGGAGTCAGATGAGACGAGCTCCTCTCTG 277
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 278 ACTGCGCGGAGGTGAAGCAATTCCGCGGAGATCAGCCGAAGAGCAAGTGGTGAA 337
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 338 ATGCTGGGAGATGGGACACCTACAAAAACAGCAAAAGCTCATAGTCAAGCTACCAG 397
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 398 GGAATTCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACATTGAGGAATC 457
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 458 AAGTGA AAAACCCCGAAGATACAGATCATGAAGAGAGGAGAGAGTCTGAA 517
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 518 CACATCCAGCAGATGACCTGAGCGTAAGTGGGACATTAGAGGCAATATTCTTCAGG 577
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 578 GATCGATACGGAACCAAGCAGCGGAACTATTACATCTCTGGCGTATGAGAGTAT 637
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
DB 638 AACCCGAGGTGGCTACTGTCAGGGACCTGAGCCATCGCCCTGTCTCTCTTAT 697
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 698 CTTCCTGAGAGATGATCTTGGGCACTGTGTGCACTGCTGCCAGAGAGGCACTCC 757
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 758 CTGACAGGATTTCAAGCCCAATATGCGGGAACGTCACAGGGGCTCAAGACCAAGAG 817
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 818 CATGTGGTAGCCACGTCACTACCAAGACCATGTGGCATCAGGACAAAGATCTATGT 877
QY 261 -GlyGlnCysSer-ProLeuGlyCysLeuIleArg-IleLeuIleAsp--GlyIleSer 278
DB 878 GGGGCAAGTGTTCCTTCTAGGCTGCTCATCCGGGATATTGATTGAATGGGATCTCTCC 937

QY 279 LeuGlyLeuThrLeuArgLeuTrp 286
DB 938 TCGGGCTCACCCCTGCGCCTCTGG 961

RESULT 9
LOCUS B1522600 919 bp mRNA linear EST 29-AUG-2001
DEFINITION 603175612F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240090 5', mRNA sequence.
ACCESSION B1522600
VERSION B1522600.1 GI:15347392
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 919)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM1605 row: 9 column: 03
High quality sequence stop: 834.
Location/Qualifiers

FEATURES
source
1. .919
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5240090"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC library."

ORIGIN
Alignment Scores: 2.42e-82 Length: 919
Pred. No.: 1335.00 Matches: 247
Score: 1335.00
Percent Similarity: 98.0% Conservative: 0
Best Local Similarity: 98.0% Mismatches: 4
Query Match: 44.8% Indels: 2
DB: 3 Gaps: 0

US-10-071-838-2 (1-549) x B1522600 (1-919)

QY 299 ProIleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCys 318
DB 1 CCGATTAACAAGATCGCCTTTAAGTTTCAGACAGACGCTCAAGAAAGCTCCAGTGT 60
QY 319 GlyProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThr 338
DB 61 GGCCCGTGGGCAAGTTTGGCAACCGGTTCTGTATACCTGGCCAGATAGAGACT 120
QY 339 ValLeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuPro 358
DB 121 GTGCTCAAGCATCTTAGGGCTCTATGAAGAACTAACAAGAAAGAGGGGACCTGCAA 180

Qy 402 TrpSerAlaSerProProArgAlaProArgSerSerThrProCys-ProGlyGlyAla-V 421
Db 719 TGGTCAGCTTCCCGCCACGGGACCTCGTTCTTCACACCCGTCTCCCTGGTGGGCTTG 778
Qy 421 alArgGluAspThrTyProValGly---ThrGlnGlyValProSerProAlaLeuAlaG 440
Db 779 TCCGGGAAGACACCTACCTCTGTGGGAACCTTCAGGGTGTGCCAGCCCC-----GGGC 832
Qy 440 InGlyGlyProGlnGlySer-----TrpArgPheLeuGlnTrpAsnSer----- 454
Db 833 CTGGGNTCAGGGAAGAACTCAGGGGTTCCTCGAAA---ATTCCCTGGCAGGGGAAC 889
Qy 455 -----MetProArgLeuProThrAsp 461
Db 890 TCCATTGCCCCCCCTCCCAACGGAA 915

RESULT 11
BM553146 1141 bp mRNA linear EST 20-FEB-2002
LOCUS BM553146
DEFINITION AGENCOURT_6542478 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742850
5', mRNA sequence.

ACCESSION BM553146
VERSION BM553146.1 GI:18791621
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1141)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsapb8-rc@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM12761 row: k column: 11
High quality sequence stop: 665.
Location/Qualifiers

FEATURES

source 1..1141
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5742850"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_11b="NIH MGC 119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC library."

ORIGIN

Alignment Scores:
Pred. No.: 2,21e-81 Length: 1141
Score: 1323.00 Matches: 276
Percent Similarity: 70.5% Conservative: 8
Best Local Similarity: 68.5% Mismatches: 51
Query Match: 44.4% Indels: 68
DB: 3 Gaps: 6

US-10-071-838-2 (1-549) x BM553146 (1-1141)

Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 102 ATGACGTTGTAGAGGTCCGGGCGAGTTGTGGGCACAAGAGCGAGAGACATCATTAAG 161
Qy 21 LysTyrgLulysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 162 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTGA 221
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 222 AGCTACACAACAACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCTCTCTG 281
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 282 ACTGCGGGGAGGCGAAGCAATTCCGCGGAGATCAGCCGAAAGACCAAGTGGTGAT 341
Qy 81 MetLeuGlyAspTrpGluLulysTyrlsSerSerArgLysLeuIleAspArgAlaTyrls 100
Db 342 ATGCTGGAGACTGGGAGAAATACAAAGCAGACGAAAGCTCATAGATCGAGCTACAA 401
Qy 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
Db 402 GGAATGCCCATGAAATCCGGGGCCGATGTGTCAAGTCCCTCTGAACACTGAGGAATG 461
Qy 121 LysLeuLysAsnProGlyArgTyrgLulileMetLysGluLysGlyLysArgSerSerGlu 140
Db 462 AAGATGAAAAACCCCGGAAGATACAGATCATGAAGAGAGGCAAGAGTCATCTGAG 521
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 522 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAGGAAGCATATATTCTTCAG 581
Qy 161 AspArgTyrgLylThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrgLulTy 180
Db 582 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGATAT 641
Qy 181 AsnProGluValGlyTyrcysArgAspLeuSerHisIleAlaIleuPheLeuLeuTy 200
Db 642 AACCCGAGGTGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 701
Qy 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlu-ArgHisSe 220
Db 702 CTTCCTGAGGAGATGCATTTCTGGCACTGTGACGCTGCTGGCCAGTGAAGAAGCATTC 761
Qy 220 rLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 762 CCTGACAGGATTTCAACAGCCCAATGGGGGACCGCCCGGGGCTCCAAGACCCACAGGA 821
Qy 240 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCy 260
Db 822 GCATGGGGGAACCCGTTACCACCCAGACCTGGGG----- 858
Qy 260 sGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGl 280
Db 858 ----- 858
Qy 280 yLeuThrLeuArgLeuTrpAspValTyrlsLeuValGluGlyGlnAlaLeuMetProIl 300
Db 858 ----- 858
Qy 300 eThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPr 320
Db 859 -----GACTCAAAACGCTCACCAGAAGACTCCCAATTGTGGCC 896
Qy 320 oTrpAlaArg---PheCysAsnArgPheValAspThr-TrpAlaArgAspGluAspThr- 338
Db 897 CCGGCGCCCTTCTTTACACGCGGTTCTTGTATCCTTGGGTCAAGCAATCGGGGACCC 956
Qy 339 --ValLeuLysHisLeuArgAlaSerMet--LysLysLeuThrArgLysGlnGlyAspL 357
Db 957 CGGTGCTCAAGCACTTTTGGGCTCTGTATATAAATAACCGGAAACGAGGGGG 1016

QY 357 euPro-----ProProAlaIySProGIuGIySer-----Sera 369
|||
|||
Db 1017 CACCCGTGTCCCCCGCCCAACCCAGCGCTTAAGAGGGGTCACATGCATCCCAACCG 1076
QY 369 laSerArgProValProAlaSerArgGIyGIyLysThrLeuCysIySGIyAspArgGlna 389
|||
|||
Db 1077 CGTGACACCCGTTTTCCTCCCGCGCGGAGAAATCACCGCATCGGCGGGGGAACAAGTA 1136
QY 389 laPro 390
|||
Db 1137 TCCCA 1141
RESULT 12
BM563571 1079 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6564823 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744077
DEFINITION 5' mRNA sequence.
ACCESSION BM563571
VERSION BM563571 GI:18810609
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1079)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-romail.nih.gov
Issue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12764 row: n column: 14
High quality sequence stop: 665.
FEATURES
source
location/Qualifiers
1.1079
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744077"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6, Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 5.38e-80 length: 1079
Score: 1302.50 Matches: 262
Percent Similarity: 91.9% Conservative: 9
Best Local Similarity: 88.8% Mismatches: 15
Query Match: 43.7% Indels: 9
DB: 3 Gaps: 1
US-10-071-838-2 (1-549) x BM563571 (1-1079)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
|||
|||
Db 100 ATGACCTGTGTAAGAGTCCGCGGTAGTTGTGTGGCACAAGACGAGACATCATATATG 159

QY 21 LysTyrGIyLysGIyHisArgAlaGlyLeuProGIuAspIySGIyProLysProPheArg 40
|||
|||
Db 160 AAATACGAAGAGGACACCGAGCTGGCTGCCAGAGACAGGGCCTTAAGCCTTTTCCA 219
QY 41 SerTyrAsnAsnAsnValAspHisLeuGIyIleValHisGluThrGIuLeuProProLeu 60
|||
|||
Db 220 AGCTACAACAACAACGTGCATCATTTGGGATTGTACATGAGACGAGCTGCCCTCTCTG 279
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
|||
|||
Db 280 ACTGCGCGGAGCGCAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 339
QY 81 MetLeuGIyAspTrpGIuLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
|||
|||
Db 340 ATGCTGGAGACTGGAGAAATACAAAGACAGAAAGCTCATAGATCGAGCGTACAG 399
QY 101 GlyMetProMetAsnIleArgGIyPrometTrpSerValLeuLeuAsnIleGIuGluMet 120
|||
|||
Db 400 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTGAACATTGAGGAATG 459
QY 121 LysLeuLysAsnProGIyArgTyrGlnIleMetLysGIyLysArgSerSerGIu 140
|||
|||
Db 460 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAGGCAAGGTCAATCTTACG 519
QY 141 HisIleGlnArgIleAspArgAspValSerGIyThrLeuArgLysHisIlePhePheArg 160
|||
|||
Db 520 CACATCCAGCGCATCGACCGGACATTAACGCGGACATTAAAGAAATATGTTCTTACGG 579
QY 161 AspArgTyrGIyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGIuGIyTyr 180
|||
|||
Db 580 GATCGATACGAACCAAGCAGCGGAACTACTCCACATCTCTTGCATATGAGAGATAT 639
QY 181 AsnProGIuValGIyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
|||
|||
Db 640 AACCCGAGGTGGCTACTGCAGGAGCACTGAGCCACATCGCGCCTTGTCTCTCTTA 699
QY 200 rLeuProGIuGIuAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGIuArgHisSe 220
|||
|||
Db 700 TTTTCTTGAGGAGATGCATTCTGGGCACTGTGCAGCTGCTGCGCAGTGAGAGGCACTC 759
QY 220 rLeuGlnGIyPheHisSerProAsn-GIyGIyThrValGlnGIyLeuGlnAspGlnGln 240
|||
|||
Db 760 CTTGACGAGATTTCACAGCCCAATGGGCGGACGTCCAGGGGGCTCCACAGACCAACAGG 819
QY 240 lu-HisValAlaAa--ThrSerGlnProLysThrMetGIyHisGln-AspLysIyAs 258
|||
|||
Db 820 AAGCATGTGGTAGCCACGTTCAACCATCAAGACCATGGGGCATCAAGGACAGAAAGA 879
QY 258 p-LeuCysGIyGlnCysSerPro-LeuGIyCysLeuIleArgIleLeu--IleAspGIyI 277
|||
|||
Db 880 TCCTATGTGGCAGTGTTCCTCCGGTTAGGGTGCCTCATCCGATATGATTGGACGGGA 939
QY 277 leSerLeuGIyLeuThrLeuArgLeuTrp 286
|||
|||
Db 940 TCTTCTTTCGGGTTCACCCCTGCGCCTGG 968
RESULT 13
BI828272 831 bp mRNA linear EST 04-OCT-2001
LOCUS 603078163F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169999 5',
DEFINITION mRNA sequence.
ACCESSION BI828272
VERSION BI828272.1 GI:15939822
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 831)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11422 row: n column: 16
High quality sequence stop: 822.

FEATURES

Source
1. 831
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5169999"
/issue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH MGC 119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 4.5e-78 Length: 831
Score: 1272.50 Matches: 261
Percent Similarity: 93.0% Conservative: 3
Best Local Similarity: 91.9% Mismatches: 10
Query Match: 42.7% Indels: 16
Gaps: 2

US-10-071-838-2 (1-549) x B1828272 (1-831)

QY 148 AspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGln 167
Db 2 GACATAAGCGGACATTAAGGAAGCATATGTTCTTCAGGGATCGATACGGAACCAAGCAG 61
QY 168 ArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyrAsnProGluValGlyTyrCys 187
Db 62 CGGAACTACTCCACATCTCTCGGCATATGAGATATAACCCGAGGTGGCTACTGC 121
QY 188 ArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAlaPhe 207
Db 122 AGGACCTGAGCCACATCGCCGCTGTCTCTCTATCTTCTGAGAGATGCAATC 181
QY 208 TrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSerPro 227
Db 182 TGCGCACTGGTGCACTGCTGGCCAGTGAAGGCACTCCCTGCAGGGATTCAAGCCCA 241
QY 228 AsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlnHisValValAlaThrSerGln 247
Db 242 AATGGCGGACCGCTCCAGGGGCTCCAGACCAACAGAGCATGTGTAGCCACGTACAA 301
QY 248 ProLysThrMetGlyHisGlnAspLysLysAspLeuCyseGlyGlnCysSerProLeuGly 267
Db 302 CCCAAGACCATGGGGCATCAGGACAGAAAGATCTATGTGGCAGTGTCCCGTTAGGC 361
QY 268 CysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuTrpAsp 287
Db 362 TGCTCATCCGATATGATTGACGGGATCTCTCTGGGCTCACCCCTGTGGAC 421
QY 288 ValTyrLeuValGluGlyGlnAlaLeuMetProIleThrArgIleAlaPheLysVal 307
Db 422 GTGTATCTGTAGAAAGCGAAGACAGGCGTTGATGCCGATAACAAGATCGCTTAAGTT 481

QY 308 GlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAsnArg 327
Db 482 CAGCAGAGCGCCTTCAAGAAAGCAGTCCAGGTGTGGCCCGTGGGCACTTTTGCAACCGG 541
QY 328 PheValAspThrTrpAla-ArgAspGluAspThrValLeuLysHisLeuArgAlaSerMet 347
Db 542 TTCGTTGATACCTGGGCCCCAGGATGAGGACACTGTGCTCAAGCATCTTAGGCCCTTAT 601
QY 347 TlyLysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGluGlnGlySe 367
Db 602 GAAGAACTAACAAGAAAGCAGGGGAGCTGCCACC-CCAGCCAAACCCGACAGGAGTC 660
QY 367 rSerAlaSerArgProValProAlaSerArgGlyGlyLysThrLeuCyseGlyAspArg 387
Db 661 GTGGCATCCAGGCGCTGTGCCGCTTACGTGGCGGAAAGACCTCTGCAAGGGGGACA 720
QY 387 GlnAlaProProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProPr 407
Db 721 GGCAGGCCCTCCAGGCCACCA-GCCCGGTT-CCGCGGCATT-TGGTCAAGCTTC-CCGCC 776
QY 407 cArgAlaProArgSerSerThrPro-----CysProGlyGlyAlaValArgG1 423
Db 777 ACGGGC-CCTCGTTCTTCCAACCTGTGTGTGGGCTGTCCG-----GA 820
QY 423 uAspThrTyr 426
Db 821 AGACACCTAC 830

RESULT 14

BI088323

LOCUS BI088323 709 bp mRNA linear EST 20-JUN-2001

DEFINITION 602852793F1 NIH_MGC_10 Homo sapiens CDNA clone IMAGE:4994398 5', mRNA sequence.

ACCESSION BI088323

VERSION BI088323.1 GI:14506653

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 709)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM11016 row: 1 column: 23

High quality sequence stop: 709.

Location/Qualifiers

1. 709

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4994398"

/cell_line="MGC36"

/lab_host="DH10B"

/clone_lib="NIH_MGC_10"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.5 kb. Library prepared by Life Technologies."

ORIGIN

Alignment Scores:

Pred. No.: 3.99e-76 Length: 709
Score: 1243.00 Matches: 235
Percent Similarity: 99.6% Conservative: 0
Best Local Similarity: 99.6% Mismatches: 1
Query Match: 41.7% Indels: 0
DB: 2 Gaps: 0

US-10-071-838-2 (1-549) x BI088323 (1-709)

QY 47 AspHisLeuGlyIleValHisGluThrGluLeuProProLeuThrAlaArgGluAlaLys 66
|||||
Db 2 GATCATTTGGGGATTGTACATGAGACGAGCTGCTCTCTGACTGCGGGAGCGGAAG 61
QY 67 GlnIleArgArgGluIleSerArgLysSerLysTrpValAspMetLeuGlyAspTrpGlu 86
|||||
Db 62 CAAATTCCGGCGGAGATCAGCCGAAAGACCAAGTGGGTGATATGCTGGAGACTGGAG 121
QY 87 LysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLysGlyMetProMetAsnIle 106
|||||
Db 122 AAATACAAAGCAGCAGAAAGCTCATAGATCAGCGCTACAAAGGAATGCCCATGAACATC 181
QY 107 ArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMetLysLeuLysAsnProGly 126
|||||
Db 182 CCGGGCCCGATGTGTAGTCTCTCTGAACACTGAGAAATGAAGTTGAACCCCGGA 241
QY 127 ArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGluHisIleGlnArgIleAsp 146
242 AGATACCAAGATCATGAAGAGAGAGGCAAGAGGTCACTTGAGCACATCCAGCGCATCGAC 301
QY 147 ArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLys 166
|||||
Db 302 CGGACGTAAGCGGACATTAAGAGACATATATCTTCAGGGATCGATACGGAACCAAG 361
QY 167 GlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluuTyrAsnProGluValGlyTyr 186
|||||
Db 362 CAGCGGAACTACTCCACATCTCTCTGCGCATATGAGAGATAAACCCGGAGTGGGCTAC 421
QY 187 CysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyrLeuProGluGluAspAla 206
|||||
Db 422 TGCAGGACCTGAGCCACATCGCCGCTGTCTCTCTCTATCTTCTTGAGGAGGATGCA 481
QY 207 PheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGlnGlyPheHisSer 226
|||||
Db 482 TTCTGGGCACTGTGTGAGCTGTGGCCAGTAGAGGCACTCCCTGAGGGAATTCCACAGC 541
QY 227 ProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGluHisValAlaIleThrSer 246
|||||
Db 542 CCAATGCGCGGAGACCGTCCAGGGGCTCCAAGACCAAGAGCATGTGTAGCCACGTCA 601
QY 247 GlnProLysThrMetGlyHisGlnAspLysLysAspLeuCysGlyGlnCysSerProLeu 266
|||||
Db 602 CAACCCAAGACCATGCGGCAATCAAGAAAGATCTATGTGGCAGTGTCCCCGTTA 661
QY 267 GlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLeuThr 282
|||||
Db 662 GGCTGCTCATCCGATATTGATTGACGGGATCTCTCTCGGGCTCACC 709

RESULT 15
BQ891586 946 bp mRNA linear EST 16-AUG-2002
LOCUS BQ891586
DEFINITION AGENCOURT_8670678 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6380487
5', mRNA sequence.

ACCESSION BQ891586
VERSION BQ891586.1 GI:22283600
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS 1 (bases 1 to 946)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsabds-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov

plate: L1CM2569 row: c column: 16
High quality sequence start: 2
High quality sequence stop: 554.

FEATURES

source

Location/Qualifiers
1..946
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6380487"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 8.84e-76 Length: 946
Score: 1240.50 Matches: 259
Percent Similarity: 85.2% Conservative: 12
Best Local Similarity: 81.4% Mismatches: 32
Query Match: 41.6% Indels: 15
DB: 5 Gaps: 4

US-10-071-838-2 (1-549) x BQ891586 (1-946)

QY 82 LeuGlyAspTrp-GluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLysGly 101
|||||
Db 4 TTGGGAGACTGGGGAGAAATACAAAGACAGAAAGCTCATAGATCGAGCGTACAAAGG 63
QY 101 yMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMetLys 121
64 AATGCCCATGAACATCCGGGGCCCGATGTGTCAAGTCTCTGAACTAGAGAAATGAA 123
Db 121 sLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGluHis 141
124 GTTGAACAAACCCCGGAAGATACCATGATCATGAAGAGAGGCGCAAGATATATTCTTCAAGGA 183
QY 141 sIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArgAs 161
184 CATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAAAGCATATATTCTTCAAGGA 243
QY 161 PArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyrAs 181
244 TCGATACGGAACCAAGACGCGGAACTACTCCATCTCTCTGCGCATATGAGAGATATAA 303
Db 181 nProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyrLe 201
304 CCCGAGGTGGGCTTACTGCAAGGACCTGAGCCACATCGCGCTTGTCTCTCTATCT 363
QY 201 uProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLe 221
Db 364 TCCTGAGGAGGATGCTTCTGGGCACTGTGTCAAGTCTGCGCAGTAGAGCACTCCCT 423
QY 221 uGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGluHis 241
Db 424 GCAGGGAATTCACAGCCCAATGGCGGAGCGTCCAGGGGCTCCAAAGCAACAGAGAGCA 483

QY	241	BValValAlaIaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCysG1	261
Db	484	TGtGTAGCCACGTACAAACCCAAAGACCATGGGGCATCAGACAAGAAAGATCTATGTGG	543
QY	261	YGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLe	281
Db	544	GCAGTGGTCCCGGTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGGCT	603
QY	281	uThrLeuArgLeuThrAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIleTh	301
Db	604	CACCTGCGCCTGTGGAGCTGATCTGGTAGAAGCGCAACAGCGCTGATGGCGATAAC	663
QY	301	rArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro-T	321
Db	664	AAGAATCGCCTTTAAGTTCAGCAGAAGCGCTCACCAAGACGGCCAAAGGGGGCCGGG	723
QY	321	rpAlaArg-PheCysAsnArg--PheValAspThrTrpAlaArgAsp---GluAspThr-	338
Db	724	GGGCAGCTTTTTCGCAACCGGGGTGCTTGATACCTGGGGGCAGGGGATGGAGCACTG	783
QY	339	--ValLeuLysHisLeu-ArgAlaSerMetLysLysLeuThrArg-----	352
Db	784	GGGGTTAAAGCCATCTTTAGGGGCGCTTATGGAGGAAAACCTAAACCAAGAAAAAGAGGG	843
QY	353	---LysGlnGlyAspLeuProProProAlaLysProGluGlnGlySerSerAlaSerArg	371
		:::	
Db	844	GGGAAACCTGGCAAAACCCCCAGCCACAACCCCCAAAGCAAAAGGGGTCGTTCCG--	900
QY	372	ProValProAlaSerArgGlyGlyLysThrLeuCysLysGly	385
		:::	
Db	901	GCAATTCCCAAGGGCGTGGGGGCGGGCGTTTCCGGCGGG	942

Search completed: April 5, 2006, 16:26:30
Job time : 6220 secs

100-443887-100

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:28:52 ; Search time 1202 Seconds
(without alignments)
3044.023 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWWAQEREDIM.....TSGPCLGHLHSSQFPFGF 549

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 segs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/abs/ABSSWEB_spool/US10071838/runat_05042006_141912_5134/app_query.fasta_1
-DB=N_Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs07
-USER=US10071838 @CCG 1_1 727 @runat_05042006_141912_5134 -NCPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_21:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	6	ABSS4706 \ Abbs4706 cDNA enco
2	2983	100.0	1964	10	ADC37382 Nuclea f Adc37382 Nuclea f
3	2978	99.8	1993	3	AAAI5001 Aaai5001 cDNA enco
4	2978	99.8	1993	13	ADT88069 Adt88069 Human pro

5	2975	99.7	1964	10	ADC37384 Adc37384 Nuclear f
6	2963	99.3	2072	5	ABX71403 Abx71403 Human cel
7	2950	98.9	2064	13	ADQ84596 Adq84596 Human tum
8	2950	98.9	2064	13	ADQ86595 Adq86595 Human tum
9	2907.5	97.5	2146	6	ABSS4707 Abbs4707 cDNA enco
10	2906.5	97.4	2304	10	ADC37380 Adc37380 Nuclea f
11	2898.5	97.2	2304	10	ADC37378 Adc37378 Nuclea f
12	2850	95.5	2084	6	ABZ11810 Abz11810 Novel hum
13	2850	95.5	2084	12	ADM44328 Adm44328 Nuclea f
14	2835	95.0	2647	10	ADC37388 Adc37388 Nuclea f
15	2827	94.8	2072	6	ABZ11809 Abz11809 Human pol
16	2827	94.8	2072	12	ADM44327 Adm44327 Novel hum
17	2827	94.8	2647	10	ADC37386 Adc37386 Nuclea f
18	2827	94.8	2647	11	ADM01991 Adm01991 Human cDN
19	2784	93.3	1752	6	ABST78751 Abst78751 DNA encod
20	2773	93.0	1862	6	ABSS4708 Abbs4708 cDNA enco
21	2191.5	73.5	7856	5	AAAS8531 Aaas8531 DNA encod
22	2191.5	73.5	7856	5	AAAS88213 Aaas88213 DNA encod
23	2191.5	73.5	7878	6	ABL64684 Abl64684 Stomach c
24	2188.5	73.4	2881	13	ACN42882 Acn42882 Human dia
25	2142.5	71.8	3799	13	ACN42884 Acn42884 Human dia
26	2140	71.7	3901	13	ACN42883 Acn42883 Human dia
27	2135.5	71.6	8435	4	AAI59633 Aai59633 Human pol
28	2135.5	71.6	8435	4	AAI59634 Aai59634 Human pol
29	2135.5	71.6	8435	10	ADC31883 Adc31883 Human nov
30	2133.5	71.5	8180	5	AAAS88212 Aaas88212 DNA encod
31	2133.5	71.5	8201	2	AAAT12170 Aaat12170 pTG4-5-CD
32	2133.5	71.5	8201	10	ADE85053 Ade85053 Farnesyl
33	2133.5	71.5	8284	5	AAAS8535 Aaas8535 DNA encod
34	2133.5	71.5	8408	4	AAI57848 Aai57848 Human pol
35	2133.5	71.5	8420	4	AAI57847 Aai57847 Human pol
36	1795.5	60.2	3396	5	AAAS8534 Aaas8534 DNA encod
37	1795.5	60.2	3904	5	AAAS88214 Aaas88214 DNA encod
38	1732	58.1	2838	5	AAAS87029 Aaas87029 DNA encod
39	1732	58.1	2838	5	AAAS84009 Aaas84009 DNA encod
40	1732	58.1	2838	5	AAAS87522 Aaas87522 DNA encod
41	1533.5	51.4	2424	5	AAAS87523 Aaas87523 DNA encod
42	1507.5	50.5	9805	4	AAAL05169 Aaal05169 Human rep
43	1507.5	50.5	9805	4	ABL98056 Abl98056 Human tes
44	1507.5	50.5	9805	4	ABL98056 Abl98056 Human tes
45	1507.5	50.5	9805	10	ADE47018 Ade47018 Human car

ALIGNMENTS

RESULT 1	
ABSS4706	
ID	ABSS4706 standard; cDNA; 1964 BP.
XX	
AC	ABSS4706;
XX	
DT	28-NOV-2002 (first entry)
XX	
DE	cDNA encoding human PRC17 protein.
XX	
KW	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;
KW	breast cancer; cytosolic; chromosome 17q11-12; gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1650
XX	/*tag= a
XX	/product= "Human PRC17 protein"
PN	WO200262958-A2.
XX	
PD	15-AUG-2002.
XX	
PF	08-FEB-2002; 2002WO-US003457.
XX	
PR	08-FEB-2001; 2001US-0267615P.
XX	

PA (TULIA-) TULARIK INC.
XX
PI LI J, Powers S, Xiang P, Peng Y;
XX
DR WPI; 2002-706902/76.
XX P-PSDB; ABG70736.
XX
PT Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PT or expression relative to normal, for example cancer.
XX
PS Claim 24; Page 62; 78pp; English.
XX
CC The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present nucleic acid sequence represents the human PRC17 gene
CC located on chromosome 17q11-12. This sequence encodes the human PRC17
CC protein of the invention
XX
SQ Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.72e-128 Length: 1964
Score: 2983.00 Matches: 549
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-071-838-2 (1-549) x ABS54706 (1-1964)

QY 1 MecAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGAGCGTGTAGAGCTCGCGGCACTGGTGGGCACAAGAGCGAGCATCATATG 60
QY 21 LysTYrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGCCTTAAGCCTTTTCA 120
QY 41 SerTYrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
Db 121 AGCTACAAACAACACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCCTCTCTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 181 ACTGCGCGGAGCGAAGCAAAATTCGCGCGGAGATCAGCCGAAAGACAAGTGGGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100
Db 241 ATGCTGGGAGACTGGGAGAAATACAAAGACAGCAAGCTCATGATCGAGCGTACAAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTCTGAACATTGAGGAATG 360
QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAAACCCCGAAGATACACAGATCATGAAGAGAGAGGCAAGGTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAAGAGCATATATTCTTCAGG 480
QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluTYr 180
Db 481 GATCGATACGAAACAAGCAGCGGGAACTACTCCACATCCTCTGGCATATGAGAGATAT 540

QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTYr 200
Db 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCCTTGTTCTCTCTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTTCTGAGAGAGATGATTCGTGGGCACTGTGTCAGCTGCTGGCCAGTGAGAGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyTYrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 661 CTGCAAGGATTTCAACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 721 CATGTGTAAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 781 GGGCAGTGTTCCTCGTTAAGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 841 CTCACCTTCGCGCTGTGGACGTGTATCTGTAGAAGCGAAACAGCGTTGATTCGAGTA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 901 ACAAGAAATCGCCTTTAAGTTTCAGACAAGCGCCTCAAGAAAGCGTCAGGTGAGCCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 961 TGGGACAGTTTTTGGCAACCGGTCGTTGATACCTGGGCCAGGATGAGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1021 AAGCATCTTAAGGGCCTCTATGAAGAACTAACAAAGAAAGAGGGGAGCCTGCCACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1081 GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCGCTGTGCCGCTTCAAGTGGCGGGAAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1141 ACCCTTCGAAGGGGAGACAGGAGCCCTCCAGGCCCCACAGCCCGGTTCCCGGGCCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1201 ATTTGGTCAGCTTCCCGCCAGCGGCACTCGTCTTCCACACCTGTCTCTGGTGGGCT 1260
QY 421 ValArgGluAspThrTYrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1261 GTCCGGGAAGACACCTACCCCTGTGGCACTCAGGGGTGCCACGCCCGCTGGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1321 GAGGACCTCAGGGTTCCTGAGATTCTGCAGTGAATCTCCATGCCCGCTCCCAACG 1380
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTYrAspPheArgGlnSerCysTrpVal 480
Db 1381 GACCTGACGTAGAGGGCCCTTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1441 CGTGCCATATCCAGAGAGCAAGCTGGCCCCCTGTGGCAGGCTGAACACCCCTCGGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyTYrThrProPheArg 520
Db 1501 CGGGTGAATCGGCTTTCGTGCAACCAAGCACTGATTCGAGCAAGGCAACCCCTTCAAGA 1560
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLysCysGlyLeuHisLys 540
Db 1561 GCTAAGGAGCAACAGCCGTGTGCTCCACCTCAGGGGCTTGGCTTGGCGGCTTCACTTG 1620
QY 541 GluSerSerGlnPheProProGlyPhe 549

Db 1621 GAAAGTCTCAGTTCCTCCAGGCTTC 1647
RESULT 2
ADC37382
ID ADC37382 standard; DNA; 1964 BP.
XX
AC ADC37382;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 215.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; dg;
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
XX
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAMI KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PSDB; ADC37383.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 215; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 1964 BP, 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.72e-128 Length: 1964
Score: 2983.00 Matches: 549
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 10 Gaps: 0

US-10-071-838-2 (1-549) x ADC37382 (1-1964)

QY 41 SerTyrAsnAsnAsnValAspHisIleuGIYIleValHisGIuThrGIuLeuProProLeu 60
Db 121 AGCTACACAACAACAGCTGCATCTTGGGAGTTGTACATGAGACGAGCTGCCTCCTCTG 180
QY 61 ThrAlaArgGIuAlaIalysGlnIleArgArgGIuIleSerArgIysSerIysTrpValAsp 80
Db 181 ACTGCGCGGAGGCGCAAGCAATTCGCGCGAGATCAGCCGAAAGACCAAGTGGTGAT 240
QY 81 MetLeuGIYAspTrpGIuIlySTYrIlysserSerArgIysLeuIleAspArgAlaTYrIlyS 100
Db 241 ATGCTGGAGACTGGGAGAAATACAAAGACAGAAAGCTCATAGATCGACGTACAAAG 300
QY 101 GIYMetProMetAsnIleArgGIYProMetTrpSerValLeuLeuAsnIleGIuGIuMet 120
Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTGAGTCCCTTGAAcATTGAGAAATG 360
QY 121 IysLeuIysAsnProGIYArgTYrGlnIleMetIysGIuIlySGIYLYsArgSerSerGIu 140
Db 361 AAGTTGAAAAACCCCGAAGATACCATGATGAAAGAGGCAAGGTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGIYThrLeuArgIysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATCGACCGGACGTAGCGGACATTAAAGAGCATATATTCTTCAGG 480
QY 161 AspArgTYrGIYThrIlySGlnArgGIuLeuLeuHisIleLeuLeuAlaTYrGIuGIuTYr 180
Db 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGATAT 540
QY 181 AsnProGIuValGIYTYrCysArgAspIleuSerHisIleAlaIleuPheLeuLeuTYr 200
Db 541 AACC CGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTTCTCCTCTAT 600
QY 201 LeuProGIuGIuAspAlaPheTrpAlaLeuValGIuLeuLeuAlaSerGIuArgHisSer 220
Db 601 CTTCCTGAGAGAGATGATTCGTGGGCACTGTGTGACGTGCTGGCAGTGAGAGGCACTCC 660
QY 221 LeuGlnGIYPhelHisSerProAsnGIYGIYThrValGIuGIuIlyLeuGIuAspGIuGlnGIu 240
Db 661 CTGCAGGATTTCAcAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 720
QY 241 HisValValAlaThrSerGlnProIlySThrMetGIYHisGlnAspIlySlyAspLeuCYs 260
Db 721 CATGTGTAGCCACGTCACAAcCCCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 780
QY 261 GIYGIuIlyCysSerProLeuGIYCYsLeuIleArgIleLeuIleAspGIYIleSerLeuGIY 280
Db 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGATCTCTCTCGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGIuGIuGlnAlaIleuMetProIle 300
Db 841 CTCACCCCTGCGCTGTGGAGCTGTACTGTGTAGAAGGCGAAcAGCGTTGATGCCGAT 900
QY 301 ThrArgIleAlaPheIysValGIuGlnIlySArgLeuThrIlySThrSerArgCYsGIYPro 320
Db 901 ACAAGAATCGCCTTTAAGTTCAcAGCAAGCGCTCAcGAAGcGTCCAGGTGTGCGCCG 960
QY 321 TrpAlaArgPheCYsAsnArgPheValAspThrTrpAlaArgAspGIuAspThrValLeu 340
Db 961 TGGCAGCTTTTTCAAcCGGTTCTGTGATACCTGGGCCAGGATGAGGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetIysIysLeuThrArgIysGlnGIYAspLeuProProPro 360
Db 1021 AAGCATCTTAGGCGCTTATGAAGAACTTAACAAGAAAGCAGGGGGAcCTGCACCCCA 1080
QY 361 AlaIysProGIuGlnGIYSerSerAlaSerArgProValProAlaSerArgGIYGIYLYs 380
Db 1081 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCGGCTTCACGTGCGGGAAG 1140
QY 381 ThrLeuCYsIysGIYAspArgGlnAlaProProGIYProProAlaArgPheProArgPro 400
Db 1141 ACCCTCTGCAAGGGGAGACAGGCAAGGCCCTCCAGGCCACAGCCCGGTTCCCGGCGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCYsProGIYGIYAla 420

Db 1201 ATTTGGTCAGCTTCCCGCCACGGGCACCTGCTTCCACACCCCTGCTGGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1261 GTCCGGGAAGACACTACCTCTGTGGGCACTCAGGGTGTGCCAGCCCGCTGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1321 GGAAGACCTCAGGGTCTCTGAGATTCTGCTGGAATCCATGCCCGCTCCCAACG 1380
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1381 GACCTGACGTAGAGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db 1441 CGTGCCATATCCACAGAGGACACAGCTGGCCCTGCTGGCAGGCTGAACACCTGCGAG 1500
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1501 CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATCCGACCAAGGCAACCCCTCAGA 1560
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db 1561 GCTAGGGACGAACAGCCGTGTCTCTCCACCTCAGGGCCTTGCTCTGCGGCTCCACTTG 1620
QY 541 GluSerSerGlnPheProGlyPhe 549
Db 1621 GAAAGTTCTCAGTCTCCTCCAGGCTTC 1647

RESULT 3
AAA15001
ID AAA15001 standard; cDNA; 1993 BP.

XX AC AAA15001;

XX DT 21-AUG-2000 (first entry)

XX DE cDNA encoding a human proliferation and apoptosis related protein.

XX KW Human; proliferation and apoptosis related protein; PROAP; psoriasis;
cell proliferative disorder; immunological disorder; hepatitis;
reproductive disorder; arteriosclerosis; cirrhosis; melanoma; lymphoma;
cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia;
asthma; diabetes mellitus; osteoarthritis; endometriosis;
uterine fibroid; menstrual cycle; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT CDS 42..1691
FT /*tag= a
FT /product= "proliferation and apoptosis related protein"

XX PN WO200023589-A2.

XX PD 27-APR-2000.

XX PF 19-OCT-1999; 99WO-US024511.

XX PR 20-OCT-1998; 98US-0172216P.

XX PR 04-FEB-1999; 99US-0118559P.

XX PR 11-FEB-1999; 99US-0172229P.

XX PR 22-APR-1999; 99US-0154336P.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lal P;

XX PI Azimzai Y, Baughn MR, Yang J, Shih LL;

XX DR WPI; 2000-339688/29.

XX DR P-PSDB; AAY84901.

XX New human proliferation and apoptosis related protein polypeptides used
PT for diagnosis, treatment and prevention of cell proliferative,
PT immunological and reproductive disorders.
XX
PS Claim 9; Page 117-118; 128pp; English.

CC The present sequence encodes a human proliferation and apoptosis related
CC protein (PROAP). The polypeptides and polynucleotides can be used for the
CC diagnosis, treatment and prevention of cell proliferative, immunological
CC and reproductive disorders. Disorders associated with decreased
CC expression or activity of include arteriosclerosis, cirrhosis, hepatitis,
CC psoriasis, melanoma, lymphoma and cancers of the breast, brain and
CC prostate, acquired immune deficiency syndrome (AIDS), allergies, anaemia,
CC asthma, diabetes mellitus, osteoarthritis, endometriosis, uterine
CC fibroids and disruptions of the menstrual cycle. Antibodies against PROAP
CC can be use in diagnosis of disorders characterized by PROAP e.g. in ELISA
CC (enzyme linked immunosorbent assays) and the polynucleotides may be used
CC to detect and quantify gene expression in biopsied tissues. These
CC techniques can also be used to monitor regulation of PROAP levels during
CC therapeutic intervention

XX SQ Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.96e-128 Length: 1993
Score: 2978.00 Matches: 548
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.8% Indels: 0
DB: 3 Gaps: 0

US-10-071-838-2 (1-549) x AAA15001 (1-1993)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20

Db 42 ATGACGTGTAGAGGTGCGCGGCAGTTGGTGCCACAAAGACGAGACATCATATG 101

QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40

Db 102 AAATACGAAGAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTCA 161

QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60

Db 162 AGCTACAACAACAAGTCGATCATTTGGGAGATTGTACATGAGACGAGCTGCTCTCTG 221

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

Db 222 ACTGCGCGGAGCGGAAGCAAAATTGCGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT 281

QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysIleLeuAspArgAlaTyrLys 100

Db 282 ATGCTGGGAGACTGGGAGAAATACAAAGCAGAAAGCTCATAGATCGAGCGTACAAAG 341

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120

Db 342 GGAATGCCCATGAAACATCCGGGGCCGATGTGTCACTCTCTGAACTGAGGAATG 401

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140

Db 402 AAGTTGAAAACCCCGGAAGATACCATCATGAAAGAGAGGCAAGAGTCATCTGAG 461

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160

Db 462 CACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGAAACATATATTCTTACG 521

QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180

Db 522 GATCGATACGAACCAAGACGCGGAATACTCCACATCTCTGCGCATATGAGAGATAT 581

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200

Db 582 AACCCGAGGTGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 641

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 642 CTTCTGAGAGAGATGCATTCTGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTTCC 701
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 702 CTGCAGGGAATTTCACAGCCCCAAATGGCGGAGCGTCCAGGGGCTCCAGAGCAACAGAGAG 761
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuGly 260
Db 762 CATGTGTAGCCACGTCAACAACCAAGACCAATGGGGCATCAGAGACAGAAAGATCTATGT 821
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 822 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 881
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 882 CTCACCCCTCGCCTGTGGGACGTGTATCTGTAGAAGCGGAACAGCGCTTGATGCCGATA 941
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 942 ACAAGATCGCCTTTAAGTTCAAGCAGAGCGCTCACGAAGACGTTCAAGTGTGGCCG 1001
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1002 TGGCAGCGTTTGTCAACCGGTTCTGTGATTAAGTGGCCAGGAGTGAAGACACTGTGCTC 1061
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1062 AAGCATCTTAGGGCTCTATGAAGAACTACAAGAAAGCAGGGGACCTGCCACCCCA 1121
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1122 GCCAAACCCGAGCAAGGTCGTCCGCATCAGGCCCTGTCCGGCTTCACTGGCGGGAAG 1181
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1182 ACCCTCTGCAAGGGGGAGCAGGCAAGGCCCTTCCAGGCCACCAAGCCCGGTTCCCGCGGCC 1241
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1242 ATTGTGTCAGCTTCCCGCCAGCGGCACTGCTTCTCCACACCTGTCTGTTGGGCT 1301
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1302 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAG 1361
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1362 GGAGGACCTCAGGGGTTCTCGAGATTCTGCACTGGAATCCATGCCCGCCTCCCAACG 1421
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1422 GACCTGAGCGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGTTC 1481
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGln 500
Db 1482 CGTGCCATATCCAGAGACAGAGCACTGGCCCTGTGGCAGGCTGAACACCCCTGCCGAG 1541
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1542 CGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGGCACCCCTTCAGA 1601
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlnGlyLeuHisLeu 540
Db 1602 GCTAGGAGACAGACCGGTGTCTCCCACTCAGGGCCTTGCTTGGCGGCTCCACTTG 1661
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1662 GAAAGTTCTCAGTTCCTCCAGGCTTC 1688

RESULT 4
ADT88069 ID ADT88069 standard; cDNA; 1993 BP.
XX AC ADT88069;
XX DT 30-DEC-2004 (first entry)
XX DE Human proliferation and apoptosis related protein (PROAP)-12 cDNA.
XX KW PROAP; proliferation and apoptosis related protein;
KW cell proliferative disorder; cancer; atherosclerosis;
KW immunological disorder; AIDS; acquired immunodeficiency syndrome;
KW allergy; reproductive disorder; infertility; gene therapy; cytostatic;
KW antiarteriosclerotic; immunosuppressive; anti-HIV; antiallergic;
KW antiinfertility; gynaecological; human; gene; BB.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
FT CDS 42..1691
FT /*tag= a
FT /product= "Proliferation and apoptosis related protein
FT (PROAP)-12"
XX PN US2004203106-A1.
XX PD 14-OCT-2004.
XX PF 05-MAY-2004; 2004US-00839882.
XX PR 19-JAN-1999; 99US-0172216P.
PR 04-FEB-1999; 99US-0118559P.
PR 11-FEB-1999; 99US-0172229P.
PR 22-APR-1999; 99US-0154336P.
PR 19-OCT-1999; 99WO-US024511.
PR 11-APR-2001; 2001US-00807452.
XX PA (INCY-) INCYTE CORP.
XX PI Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lal P;
PI Azimzai Y, Baughn MR, Yang J, Shih LL;
XX DR WPI; 2004-728011/71.
DR P-PSDB; ADT88050.
XX PT New human proliferation and apoptosis related proteins and
PT polynucleotides for diagnosing, preventing or treating disorders
PT associated with aberrant protein expression, e.g. cancer, AIDS,
PT atherosclerosis or infertility.
XX PS Claim 5; SEQ ID NO 31; 85pp; English.
XX SQ Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.96e-128 Length: 1993
Score: 2978.00 Matches: 548
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.8% Indels: 0
DB: 13 Gaps: 0

US-10-071-838-2 (1-549) x ADT88069 (1-1993)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 42 ATGAGCGTGTAGAGTCCGGGCGAGTTGGTGGGCACAAGACGAGAGCATCATATATG 101

QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 102 AAATAACGAAAGGAGACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCGA 161

QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 162 AGCTACAAACAACACGTGATCATTTGGGATTTGTACATGAGACGAGCTGCCCTCTCTG 221

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 222 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 281

QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 282 ATGCTGGAGACTGGGAGAAATACAAAGCAGAGAAAGCTCATAGTCAGCGTACAG 341

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 342 GGAATGCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTCTGAACACTGAGGAAATG 401

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 402 AAGTTGAAAAACCCGGAAGATACCATGATCAAGAGAGAGGCGCAAGGTCTATCTGAG 461

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 462 CACATCCAGCGCATGCAACGGGACGTAAAGCGGACATTAAAGAACATATATCTTCAGG 521

QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 522 GATCGATACGGAACCAAGCAGCGGAACACTCTCCACATCTCTCGGCATATGAGAGAT 581

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 582 AACCCGGAGGTGGGCTACTGCAAGGACCTGACCATCGCCGCTTCTCTCTCTAT 641

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 642 CTTCTGAGGAGGATGCATTCTGGGCATGTGTGCAAGCTGTGCGCCAGTGAAGGCACTCC 701

QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 702 CTGACGGGATTTCAGACCCAAATGGCGGACCTCCAGGGGCTCCAAGACCAACAGAGAG 761

QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 762 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGACAGAAGAAATCTATGT 821

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 822 GGGCAGTGTCCCGCTTAGGCTGCCTCATCCGATATTGATTGACGGGATCTCTCGGG 881

QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 882 CTCACCCCTGCGCTGTGGGACGTGATCTGTGAAGGCGAACAGCGCTTGATGCCGATA 941

QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 942 ACAAGAAATCGCCTTTAAAGTTTCAGCAGAAAGCGCTCACGAAGACGTCCAGGTGCGCCG 1001

QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1002 TGGGCACGTTTGTGCAACCGGTTGTTGATACCTGGGCCAGGAGTGAAGACACTGTGCTC 1061

QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1062 AAGCATCTTAGGGCCTCTATGAAGAAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1121

Db 1062 AAGCATCTTAGGGCCTCTATGAAGAAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1121

QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1122 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCCCTGTGCGGCTTCACGTGCGGGAGAG 1181

QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1182 ACCCTCTGAAGGGGACAGGACAGGCCCTCCAGGCCCAAGCCCGGTTCCCGCGGCC 1241

QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1242 ATTTGTCAGCTTCCCGGCCACGGGCACTCGTCTTCACACACCTGTCTGTGGGGCT 1301

QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1302 GTCCGGGAGACACCTACCTGTTGGGCACTCAGGGGTGTGCCAGCCCGGCTGCTCAG 1361

QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1362 GAGAGACCTCAGGGTCTCTGAGATTCTGCAGTGAATTCATGCCCGCTCCCAACG 1421

QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1422 GACCTGACGTAGAAGGCCCTTGTTCCGCATTATGATTCAGACAGAGCTGTGGGTC 1481

QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1482 CGTGCATATCCCAAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTGGCAG 1541

QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1542 CGGGTAGATGCGCTTTCGCTGCACCAACACTGATTCGACCAAGGACACCCCTTCAGA 1601

QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1602 GCTAGGAGACCAACGCCGTGTGCTCCACCTCAGGGCCTTGCTCTGCGGCTCCACTTG 1661

QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1662 GAAAGTTCTCAGTTCCTCTCAGGCTTC 1688

RESULT 5
ADC37384
ID ADC37384 standard; DNA; 1964 BP.
XX
AC ADC37384;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 217.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASahi KASEI KK.
XX

PI Matbuda A, Muramatsu S;
XX WPI; 2003-505282/47.
DR P-PSDB; ADC37385.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 217; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
XX
SQ Sequence 1964 BP; 473 A; 552 C; 561 G; 378 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4e-128 Length: 1964
Score: 2975.00 Matches: 548
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: 10 Gaps: 0

US-10-071-838-2 (1-549) x ADC37384 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 1 ATGACGTGGTAGAGTTCGCGGCAGTGTGGGCACAAGAGCGAGAGACATCATTTATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 61 AAATACGAAAAGGAGACACCGAGCTGGGTGCCAGAGGACAAAGGGCCTTAAGCCTTTTCA 120
QY 41 SerTyrAsnAsnAsnValAspHisIleLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 121 AGCTACAAACAACAGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCTCTCTTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 181 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAGAGCAAGTGCGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 241 ATGCTGGAGACTGGGAGAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACATTGAGGAAATG 360
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 361 AAGTTGAACAAACCCCGAAGATACCAATCATGAAGGAGGCAAGGTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 421 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGACATATATTCTTCAG 480
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 481 GATCGATACGGAACAAGCAGCGGAATCACTCCACATCCTCGGCATATGAGAGATAT 540
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
DB 541 AACCCGAGGTGGCTACTGCAAGGAGCTGAGCCACATCGCCCTTGTCTCTCTTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

DB 601 CTTCCTGAGAGGATGCATTCTGGGCACGTGTGACAGCTGTGCGCCAGTGAGAGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 661 CTGCAGGATTTCAACAGCCCAATGGGGGACCGTCCAGGGGGTCCAAAGACCAACAGAG 720
QY 241 HisValAlaIleThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
DB 721 CATGTGTAGCCACGTCAACACCCACAGCATTGGGGCATCGAGACAAGAAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTTCGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 841 CTCACCTGCGCTGTGGACGTGTATCTGTGTAAGAGGCAAGCGCTTGATGCCGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 901 ACAAGAATCGCCTTTAAGTTACAGAGAAGCGCTCACGAAGACGTCCAGGTGTGGCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 961 TGGGCACGTTTTTGCACACCGGTTCTGTGATACCTGGCCAGGATGAGACACTGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1021 AAGCATCTTAGGGCTCTATGAAGAAACTTAACAAGAAAGCAGGGGGACCTGCAACCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTACGTTGCGGGAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1141 ACCCTCTGCAAGGGGAGCAGCAGGCCCTCCAGGCCACCGCCGGTTCCCGCGGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1201 ATTTGTACGCTTCCCGGCCACGGGACCTCGTCTTCCACACCTGTCTGTGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCCGGAGACACCTACCTCTGTGGCACTCAGGGTGTGCCAGCCCGCCTGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1321 GGAGACCTCAAGGTTCTTGAGATTCTCGACGTGAATTCATGCCCGCTCCCAAG 1380
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1381 GACCTGAGCTAGAGGGCCCTGTGTCGCGCATATGATTTCAGACAGAGCTGTGGGTG 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
DB 1441 CGTGCCATATCCAGAGGACCAAGCTGCCCCCTGCGCAGGCTGAACACCTGCGAG 1500
QY 501 ArgValArgSerAlaPheAlaIleProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1501 CGGGTAGATCGGCTTTCGCTGCAACCCAGCACTGATTCCGACCAAGGGCAACCCCTTCA 1560
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLysLeuCysGlyLysHisLeu 540
DB 1561 GCTAGGAGCAACAGCGGTGTCTCCACCTCAGGGGCTTGTGCGGCTTCACTTG 1620
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1621 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1647
RESULT 6
ABX71403
ID ABX71403 standard; cDNA; 2072 BP.

XX ABX71403;
 AC
 XX 14-APR-2003 (first entry)
 DT
 XX Human cell cycle-associated cDNA from clone DKFZphes3_35p22.
 DE
 XX Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
 KM
 XX Homo sapiens.
 OS
 PN WO200112659-A2.
 XX
 PD 22-FEB-2001.
 XX
 PF 18-AUG-2000; 2000WO-IB001496.
 XX
 PR 18-AUG-1999; 99US-0149499P.
 XX 28-SEP-1999; 99US-0156503P.
 XX
 PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
 PI Wiemann S;
 XX
 XX WPI; 2001-327840/34.
 XX P-PSDB; ABUS3234.
 DR
 XX
 PT Nucleic acids having the sequences of clones isolated from libraries of
 PT different human tissues, useful in recombinant DNA methodologies.
 XX
 XX Claim 1; Page 867; 1095pp; English.
 XX
 CC This invention describes novel polynucleotides and polypeptides isolated
 CC from human cDNA libraries which can be used for gene therapy or in
 CC vaccines. The polynucleotides of the invention and antibodies encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The products of the
 CC invention may also be used to identify modulators of expression and
 CC activity and to down regulate expression and activity. The antibodies of
 CC the invention may also be used as diagnostic agents for detecting the
 CC presence of polypeptides in samples. This sequence encodes a polypeptide
 CC described in the disclosure of the invention
 CC
 XX
 SQ Sequence 2072 BP; 510 A; 579 C; 586 G; 397 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.5e-127 Length: 2072
 Score: 2963.00 Matches: 545
 Percent Similarity: 99.6% Conservative: 2
 Best Local Similarity: 99.3% Mismatches: 2
 Query Match: 99.3% Indels: 0
 DB: 5 Gaps: 0
 US-10-071-838-2 (1-549) x ABX71403 (1-2072)
 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
 99 ATGACCTGTGTAGAGCTCGCGGCAGTGTGGGCACAAGACGAGAGACATCATATG 158
 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 159 AAATACGAAAGGGAACCGAGCTGGCTGCCAGAGCAAGGGGCTTAAGCCTTTTCA 218
 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
 219 AGCTACAACACCAACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 278
 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
 279 ACTGCGCGGAGCGAAGCAAAATTCGCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 338
 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100

Db	339	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAACCTCATAGATCGAGCGGTACAG	398
QY	101	GLYMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet	120
Db	399	GGAATGCCATGAACATCCGGGGCCGATGTGTGTCACTCTCTGAACACTGAGGAAATG	458
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
Db	459	AAGTTGAAAAACCCCGAAGATACAGATCATGAAAGAAAGGGCAAGATCATCTGAG	518
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
Db	519	CACATCCAGCGCATCGACCCGGAGCGTAAGCGGACATTAAGGAAGCATATATCTTCAGG	578
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
Db	579	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGGCATATGAGAGTAC	638
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIlePheLeuLeuTyr	200
Db	639	AACCCGGAGGTGGCTACTGTCAGGGACTGAGCCACATCGCCGCTTGTTCTCTCTAT	698
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	699	CTTCCTGAGAGGATGATCATCTTGGGCACTGGTGACGTGCTGGCCAGTAGAGCACTCC	758
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	759	CTGCAGGGAATTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG	818
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	819	CATGTGTAGCCACAGTCACACCAAGCAAGACCATGGGGCATCAGAGACAAGAAAGATCATGT	878
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	879	GGGCAGTGTTCCCGTTAGGCTGCTCATCCGATATTGATTGACCGGATCTCTCTCCGG	938
QY	281	LeuThrLeuArgLeuLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle	300
Db	939	CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAACAAGCGCTGATGCCGATA	998
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	999	ACAAGAAATCGCCTTTAAGTTTCAGCAGAAGCGCTCAGAAAGACGTCCAGTGTGCGCG	1058
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1059	TGGGCACGTTTTTGCAACCGGTTCTGTATACCTGGCGCAGGATGAGACACTGTGCTC	1118
QY	341	LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1119	AAGCATTTAAGGCGCTCTATGAAAGAACTAACAAAGAAAGGGGAGCCTGCCACCCCCA	1178
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys	380
Db	1179	GCCAAACCCGAGCAAGGGGTCTGCGCATCCAGGCGCTGTGCGGCTTCACGTGGCGGAAAG	1238
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1239	ACCCTCTGCAAGGGGACAGGCGCCCTCCAGGCCCCACAGCCCGGTTCCCGGCGCC	1298
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla	420
Db	1299	ATTGTGTACGTTCCCGCGCAAGGGAACCTCGTTCTTCCACACCTGTCTGTGTGGGCT	1358
QY	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1359	GTCCGGGAAGACACCTACCTCTGTGGGCACTCAGGGGTGTGCCAAGCCCGGCTGTCTAG	1418
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnIleTrpAsnSerMetProArgLeuProThr	460
Db	1419	GGAGAGCTCAGGGTTCCTGTGAGATTCTGTCAGTGGAATCCATGCCCCGCTCCCAAG	1478

QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1479 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGCTGGTTC 1538
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1539 CGTGCCATATCCAGAGAGACAGCAGCTGGCCCCCTGCTGCAGGCTGAACACCCCTGCGAG 1598
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1599 CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1658
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyGlyLeuHisLeu 540
DB 1659 GCTAGGAGACGACAGCAGTGTGCTCCCACTCAGGGCCTTGCTGCGGCTCCACTTG 1718
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1719 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1745
RESULT 7
ADQ84596
ID ADQ84596 standard; cDNA; 2064 BP.
XX
AC ADQ84596;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1410.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
PS Claim 1; SEQ ID NO 1410; 5504bp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)

CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.91e-127 Length: 2064
Score: 2950.00 Matches: 547
Percent Similarity: 99.3% Conservative: 0
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 98.9% Indels: 2
DB: 13 Gaps: 0
US-10-071-838-2 (1-549) x ADQ84596 (1-2064)
QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 99 ATGACGTGTAGAGGTGCGGCGCAGTTGTGTGGCACAAGACGAGACATCATTTATG 158
QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 159 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGGCAAGGGCCTTAAGCCTTTTCA 218
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHis--GluThrGluLeuProProL 60
DB 219 AGCTACACACAACAGTCGATCATTTGGGATTGTACAGTGAAGACGAGCTGCCTC 278
QY 60 eutThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValA 80
DB 279 TGACTGCGCGGAGGCGAAGCAATTCCGGCGGAGATCAAGCAAGCAAGTGGGTG 338
QY 80 bPmetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrL 100
DB 339 ATATGCTGGAGAGCTGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCTTACA 398
QY 100 YsGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIuM 120
DB 399 AGGGAATGCCCATGAACATCCGGGCCCGATGTGTCACTCTCTGAACACTGAGGAAA 458
QY 120 etLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerg 140
DB 459 TGAAGTGAATAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCATCTG 518
QY 140 IuHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhea 160
DB 519 AGCACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAGAGCATATATTCTTCA 578
QY 160 rgaAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluT 180
DB 579 GGGATCGATACGGAACCAAGCAGCGGGAATACTCCACATCTCTGCGCATATGAGGAGT 638
QY 180 YrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuT 200

Db 639 ATAAccCGAGGTGGGCTACTGCAGGAGCCTGAGCCACATCGCCGCTTGTCTCTCT 698
QY 200 YrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 220
Db 699 ATCTTCTGAGGAGGATGCATTTCTGGGCACTGTGCAGCTGTGGCCAGTGAGAGGCACT 758
QY 220 erLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 759 CCCTGCAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAG 818
QY 240 LuHisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeu 260
Db 819 AGCATGTGTAGCCACGTCACAACCCAGACCATGGGGCATCAGGACAAGAAATCTAT 878
QY 260 YsGlyGlnCySerProLeuGlyCySLeuIleArgIleLeuIleAspGlyIleSerLeu 280
Db 879 GTGGGCAGTGTTCCCGTTAGGCTGCTCATCCGGATATTGATTGACGGGATCTCTCTG 938
QY 280 lYLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetPro 300
Db 939 GGCTCACCCTGCGCTGTGGGACGTGATCTGTAGAAGCGCAAGCAGCGTTGATGCCGA 998
QY 300 lEThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCySgLYP 320
Db 999 TAACAAGAAATCGCCTTAAAGTTCAAGCAGAAAGCGCTCAAGAAAGCTCCAGGTGTGCC 1058
QY 320 roTrpAlaArgPheCyAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 340
Db 1059 CGTGGGCACTTTTGCAACCGGTCGTGTGATACCTGGGCCAGGAGTGAAGACACTGTGC 1118
QY 340 euLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1119 TCAAGCATCTTAGGCTCTATGAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCC 1178
QY 360 roAlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyYL 380
Db 1179 CAGCCAAACCCGAGCAAGGGTCGTGGCATCAGGCCCTGTGCCGCTTCACTGGCGGGA 1238
QY 380 YsThrLeuCySlySgLYAspArgGlnAlaProProGlyProProAlaArgPheProArgP 400
Db 1239 AGACCCTCTGCAAGGGGGAACAGGACGGCCCTCCAGGCCACACCGCGGTTCCCGCGC 1298
QY 400 roIleTrpSerAlaSerProProArgAlaProArgSerSerThrProCySProGlyGlyA 420
Db 1299 CCATTTGGTCAGCTTCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGG 1358
QY 420 lAValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaG 440
Db 1359 CTGTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCCAAGCCCGGCTGGCTC 1418
QY 440 lnglyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProT 460
Db 1419 AGGGAGGACCTCAGGGTCTCTGAGATTCTTGCAGTGGAATCCATGCCCGCTCCCAA 1478
QY 460 hrAspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCySTpV 480
Db 1479 CGGACCTGGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGAGCTGTGGG 1538
QY 480 aLArgAlaIleSerGlnGluAspGlnLeuAlaProCySTpGlnAlaGluHisProAlaG 500
Db 1539 TCCGTGCATATCCAGAGGAGACCAAGCTGGCCCTGTGCAGGCTGAACACCCCTGCGG 1598
QY 500 lAArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheA 520
Db 1599 AGCGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATCCGACAGGGGACCCCTTCA 1658
QY 520 rGAlaArgAspGlnGlnProCySAlaProThrSerGlyProCySLeuCySgLYLeuHisL 540
Db 1659 GAGCTAGGGACGAACAGCCGTGTGCTCCCACTCAAGGCCCTTGCTTGGCGGCTCCACT 1718
QY 540 euGluSerSerGlnPheProProGlyPhe 549
Db 1719 TGAAGAATTCTCAGTTCCCTCCAGGCTTC 1747

RESULT 8
ADQ86595
ID ADQ86595 standard; cDNA; 2064 BP.
XX
AC ADQ86595;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #3468.
XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 3468; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector;
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a

CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.

XX Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5.91e-127	Length:	2064
Score:	2950.00	Matches:	547
Percent Similarity:	99.3%	Conservative:	0
Best Local Similarity:	99.3%	Mismatches:	2
Query Match:	98.9%	Indels:	2
DB:	13	Gaps:	0

US-10-071-838-2 (1-549) x ADQ86595 (1-2064)

QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
DB	99	ATGACGTGTAGAGGTCCGGGCACTTGGTGGCCACAAGACGAGAGACATCATTTATG	158
QY	21	LysTyrGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
DB	159	AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTTAAGCCTTTTCA	218
QY	41	SerTyrAsnAsnValAspHisLeuGlyIleValHis--GluThrGluLeuProProL	60
DB	219	AGCTACAACAAACAGTCGATCATTTGGGGATTGTACAAGTGAGACGAGCTGCCTCTC	278
QY	60	eUThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValA	80
DB	279	TGACTGCGCGGAGCGAAGCAATTCGGGGGAGATCAGCCGAAAGACAAAGTGGTGG	338
QY	80	spMetLeuGlyAspTrpGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrL	100
DB	339	ATATGCTGGAGACTGGGAGAATACAAAAGCAGACAGAAAGCTCATAGATCGAGCTACA	398
QY	100	ysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluM	120
DB	399	AGGGAATGCCCATGAACATCCGGGGGCCCATGTGTCTCCTCTGAACACTGAGGAAA	458
QY	120	etLysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerg	140
DB	459	TGAAGTTGAAAAACCCCGAAGATACCATGATGAGAGAGGCAAGAGTCACTCTG	518
QY	140	IuHisIleGlnArgIleAspArgAspValSergLysThrLeuArgLysHisIlePhePheA	160
DB	519	AGCACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGAACATATATCTTCA	578
QY	160	rgAspArgTyrGlyThrLysGlnArgGlnLeuLeuHisIleLeuLeuAlaTyrGlnGluT	180
DB	579	GGGATCGATACGGAACCAAGCAGCGGGAATACTCCACATCTCTCGCATATGAGGAGT	638
QY	180	yrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuT	200
DB	639	ATAACCCGAGGTGGGCTACTGCAAGGACTGAGCCACATCGCGCTTGTCTCTCTCT	698
QY	200	yrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSergLysArgHisS	220
DB	699	ATCTTCTGAGGAGATGATCTGTGGGCACTGTGCACTGTGCTGCGCAGTGAGAGCACT	758
QY	220	erLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnG	240
DB	759	CCCTGACAGGATTTCACAGCCCAATGGGGGACCGTCCAGGGGCTCCAAGACCAAGG	818
QY	240	IuHisValValAlaThrSergLysProLysThrMetGlyHisGlnAspLysLysAspLeuC	260
DB	819	AGCATGTGTAGCCACGTACACCAACCAAGACCATGGGGCATCAGGACCAAGAAAGATCTAT	878
QY	260	ysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSergLeuG	280
DB	879	GTGGGCAGTGTTCCTCCGTTAGGCTGCTCTCATCCGATATTGATTGACGGGATCTCTCTG	938
QY	280	lyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProI	300

DB	939	GGCTCACCTCGCCTGTGGGACGTGTATCTGTAGAAGGCGAACAGCGTTGATGCCGA	998
QY	300	leThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSeraArgCysGlyP	320
DB	999	TAAACAAGATCGCCTTTAAGTTCAAGCAGAACGCCCTCACGACAGCTCCAGGTGTGGCC	1058
QY	320	roTPRAlaArgPheCysAsnArgPheValAspThrTPRAlaArgAspGluAspThrValL	340
DB	1059	CGTGGGCACTTTTTCACAACCGGTTCTGTGATACCTGGGCCAGGAGATGAGACACTGTGC	1118
QY	340	eulYSHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProp	360
DB	1119	TCAAGCATCTTAGGGCTCTATGAAGAACTTAACAAAGAAAGGAGGAGCTGCACACCC	1178
QY	360	roAlaLysProGlnGlnGlySerSeraLysSerArgProValProAlaSerArgGlyGlyL	380
DB	1179	CAGCCAAACCGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCACGTGCGGGA	1238
QY	380	ysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgP	400
DB	1239	AGACCTCTGCAGAGGGGACAGGCAAGGCCCTCCAGGCCACAGCCCGGTTCGCCGGC	1298
QY	400	roIleTrpSeraLysSerProProArgAlaProArgSerSerThrProCysProGlyGlyA	420
DB	1299	CCATTGTGTACGTTTCCCGCCACGCGGCACTCTCTTCCACACCTCTCTGTGTGGG	1358
QY	420	laValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaG	440
DB	1359	CTGTCCGGGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTC	1418
QY	440	lnglyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProT	460
DB	1419	AGGAGACCTCAAGGTTCTCTGAGATTCCTGACAGTGAATCCATGCCCGCTCCCA	1478
QY	460	hrAspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpY	480
DB	1479	CGGACCTGAGAGTGAAGGGCCCTGTGCTCCGCATTATGATTCAAGACAGAGCTGTGG	1538
QY	480	alArgAlaIleSergLysGlnLysAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaG	500
DB	1539	TCCGTGCCATATCCCAAGGAGGACCACTGGCCCCCTGTGGCAGGCTGAACACCTTGG	1598
QY	500	IuArgValArgSeraLysPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheA	520
DB	1599	AGCGGTGAGATCGGCTTTCGTGCAACCCAGCACATGATTCGACAGGCGACCCCTTCA	1658
QY	520	rgAlaArgAspGlnGlnProCysAlaProThrSergLysProCysLeuCysGlyLeuHisL	540
DB	1659	GAGCTAGGAGCAACAGCCGTGTGTCCCACTCAGGGGCTGTGCTGTGCGGCTGCACCT	1718
QY	540	eulYSerSergLysPheProProGlyPhe	549
DB	1719	TGGAAGTTCTCAGTTCCTCCCTCCAGGCTTC	1747

RESULT 9
 ABS54707 standard; cDNA; 2146 BP.
 ID ABS54707
 AC ABS54707;
 DT 28-NOV-2002 (first entry)
 XX
 DE cDNA encoding human PRC17 protein splice variant 1.
 XX Human; PRC17; prostate cancer; ovarian cancer; lung cancer;
 KW breast cancer; cytostatic; chromosome 17q11-12; gene; ss; splice variant.
 XX Homo sapiens.
 OS
 XX
 PH Key
 FT CDS
 FT 1..1832
 FT /*tag= a
 FT /product= "Human PRC17 protein splice variant 1"

FT /transl except= (pos:421..422, aa:Asp)
FT /note= "This codon has an apparent 1 nucleotide deletion
FT which alters the reading frame"
XX
PN WO200262958-A2.
XX
PD 15-AUG-2002.
XX
PF 08-FEB-2002; 2002WO-US003457.
XX
PR 08-FEB-2001; 2001US-0267615P.
XX
PA (TULIA-) TULARIK INC.
XX
PI Li J, Powers S, Xiang P, Peng Y;
XX
DR WPI; 2002-706902/76.
DR P-PSDB; ABG70737.
XX
PT Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PT or expression relative to normal, for example cancer.
XX
PS Claim 24; Page 63-64; 78pp; English.
XX
CC The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present nucleic acid sequence encodes the human PRC17 protein
CC splice variant 1 of the invention. The human PRC17 gene is located on
CC chromosome 17q11-12
XX
SQ Sequence 2146 BP; 494 A; 619 C; 616 G; 417 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.49e-125 Length: 2146
Score: 2907.50 Matches: 547
Percent Similarity: 89.7% Conservative: 0
Best Local Similarity: 89.7% Mismatches: 2
Query Match: 97.5% Indels: 62
DB: Gaps: 1

US-10-071-838-2 (1-549) x ABS54707 (1-2146)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db |||||
QY 1 ATGACGCTGTAGAGTCGCGGCGAGTGTGTGGGCCAACAGAGCGAGGACATCATATATG 60
Db |||||
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db |||||
QY 61 AAATAACGAAAGGAGCACCGAGCTGGGCTGCCAGAGACAGGAGGCTTAAGCCTTTTCA 120
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHis----- 53
Db |||||
QY 121 AGCTACAAACAACACCTGCATCTTTGGGGATTGTACAGTCCCTGCCCTCTGGAGTCA 180
QY 53 ----- 53
Db 181 GCCCCACAGGAAGCCCTGTCTCTCCCTGTGCTTCTCTGGGCTGAGCCCTGAG 240
QY 53 ----- 53
Db 241 CTGAAAAGGACAGAGCCAGTCTTTCTGGGGGTCCGGACCCAGGCTGGGGCGCTCCAG 300
QY 54 -----GluThrGluLeuProPro 59
Db 301 GCCCCGTGACGTTCTCAGCTCTGCGGTGGTCCCTTACAGTGAAGAGGAGCTGCTCTCT 360

QY 60 LeuThrAlaArgGluValAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpVal 79
Db |||||
QY 361 CTGACTGCGCGGAGAGCGAAGCAAAATTCGGCGGAGATCAAGCCGAAAGCAAGTGGGT- 419
QY 80 AspMetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyr 99
Db |||||
QY 420 GATATGCTGGGAGACTGGGAGAAATACAAAAGCAGACAGAAAGCTCATAGATCGAGCTAC 479
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlu 119
Db |||||
QY 480 AAGGAATGCCCATGAACATCCGGGGCCGATGTGTCAGTCTCTGAAACATTGAGGAA 539
QY 120 MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysArgSerSer 139
Db |||||
QY 540 ATGAAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGGAGAAAGGCAAGGTCTCT 599
QY 140 GluHisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 159
Db |||||
QY 600 GAGCACATCCAGCGCATCGACCGGGACGTAAGCGGACATTAAGGAAGCATATATTCTTC 659
QY 160 ArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGlu 179
Db |||||
QY 660 AGGATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTCTGGCATATGAGAG 719
QY 180 TyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeu 199
Db |||||
QY 720 TATAACCCGGAGGTGGCTACTGACGGAACCTGAGCCACATCGCGCTGTCTCTCTC 779
QY 200 TyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 219
Db |||||
QY 780 TATCTTCTGAGAGAGATGATCTTGGGCACTGTGTACGCTCTGGCCAGTGAGAGGCAC 839
QY 220 SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGln 239
Db |||||
QY 840 TCCCTGCAGGGATTTCACAGCCCAATGGCGGAGCCGTCCAAGGGGCTCCAAGACCAAG 899
QY 240 GluHisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeu 259
Db |||||
QY 900 GAGCATGTGTAGCCACGTCACAAACCAAGACCATGGGCTTCAGACAAAGAAATCTA 959
QY 260 CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu 279
Db |||||
QY 960 TGTGGCAGTGTCCCGCTTAGGCTCCTCATCCGATATGATTGACGGGATCTCTCTC 1019
QY 280 GlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetPro 299
Db |||||
QY 1020 GGGCTCACCCGCGCCTGTGGGACGTATCTGTAGAGCGGAACAGCGCTGATGCCG 1079
QY 300 IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGly 319
Db |||||
QY 1080 ATAACAAGAAATCGCCTTTAAGTTACAGAGAAGCGCCTCACGAAGAAGCTCCAGGTGGC 1139
QY 320 ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 339
Db |||||
QY 1140 CCGTGGCAGCTTTTGTCAACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTG 1199
QY 340 LeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 359
Db |||||
QY 1200 CTCAAAGCATTTAAGGCTCTATAGAAATAACAAGAAAGCAGGGGACCTGCCACC 1259
QY 360 ProAlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGly 379
Db |||||
QY 1260 CCAAGCAAAACCGAAGGAGGTCTCGGCATCCAGGCTGTGCCGCTTACGTTGGCGG 1319
QY 380 LysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg 399
Db |||||
QY 1320 AAGACCTCTGCAAGGGGACAGGAGGCCCTCCAGGCCCAACAGCCGCTTCCGCGG 1379
QY 400 ProIleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGly 419
Db |||||
QY 1380 CCCATTGTGTCAGCTTCCCGCCAGCGGGACCTCGTTCTTCCACACACCTGTCTGTGGG 1439
QY 420 AlaValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAla 439

DB 1440 GCTGTCGGGAAGACACCTACCTGTGGCACTCAGGGTGTGCCAGCCCGGCTGGCT 1499
QY 440 GlnGlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuPro 459
DB 1500 CAGGAGAGACCTCAGGGTTCCTGAGATTCTCTGACGTGAACCTCATGCCCCGCTCCCA 1559
QY 460 ThrAspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrp 479
DB 1560 ACGGACCTGGACGTAGAGGGCCCTGTGGTCCGCATTATGATTTCAGACAGAGCTGCTGG 1619
QY 480 ValArgAlaIleSerGlnIleAspGlnLeuAlaProCysTrpGlnAlaGluHisProAla 499
DB 1620 GTCCGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGGCG 1679
QY 500 GluArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPhe 519
DB 1680 GAGCGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGGCACCCCTTC 1739
QY 520 ArgAlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHis 539
DB 1740 AGAGCTAGGGAGCAACACAGCCGTGTGCTCCACCTCAGGGCCTTGCTGTGGGCTCCAC 1799
QY 540 LeuGlnSerSerGlnPheProGlyPhe 549
DB 1800 TTGGAAGTTCTCAGTTCCTCCCAAGGCTTC 1829

RESULT 10

ADC37380
ID ADC37380 standard; DNA; 2304 BP.

AC ADC37380;

DT 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 213.

KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;

KW cancer; infectious disease; bone disease; AIDS;

KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;

KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;

KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

OS Homo sapiens.

PN WO2003048202-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.

PR 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335839P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

PA (ASAH) ASAMI KASEI KK.

PI Matsuda A, Muramatsu S;

DR WPI; 2003-505282/47.

DR P-PSDB; ADC37381.

PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.

PS Claim 4; SEQ ID NO 213; 938bp; English.

CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating

CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.

XX Sequence 2304 BP; 569 A; 610 C; 636 G; 489 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.52e-125	Length:	2304
Score:	2906.50	Matches:	540
Percent Similarity:	98.4%	Conservative:	0
Best Local Similarity:	98.4%	Mismatches:	2
Query Match:	97.4%	Indels:	8
DB:	10	Gaps:	1

US-10-071-838-2 (1-549) x ADC37380 (1-2304)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 362 ATGACGTGGTAGAGGTCGGGGCAGTTGTGGGCACAAGACGAGACATCATATG 421
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 422 AATATCGAAAGGACACCGAAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCTTTTCCA 481
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 482 AGCTACAAACAACAACGTCGATCATTTGGGATTGTATGAGACGAGCTGCTCTCTG 541
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 542 ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCACCGAAAGAGCAAGTGGGTGAT 601
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 602 ATGCTGGGAGACTGGGAGAAATACAAAGACAGCAAGCTCATAGATGAGCGTACAAG 661
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 662 GGAATGCCCATGACAATCCGGGGCCCGATGTGTCACTCTCTGAACACTGAGGAATG 721
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 722 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGAGGAGCAAGGTCATCTGAG 781
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 782 CACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAAGAACATATATTCTTCAGG 841
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 842 GATCGATACGGAACCAAGCAGCGGAACCTACTCCATCTCTGGCATATGAGAGTAT 901
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleAspPheLeuLeuTyr 200
DB 902 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTCTCTCTAT 961
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 962 CTTCTGAGGAGGATGCAATTCGGGCACGTGTGCAGCTGTGGCCAGTGAAGGCACTCC 1021
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 1022 CTGCAAGGATTTCACAGCCCAATGCGGACCGTCCAGGGGCTCCAAAGACCAAGAGAG 1081
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 1082 CATGTGTAGCCACGTCACAACCCAGAGCAATGGGGCATCAGAGCAAGAAAGATCTATGT 1141
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 1142 GGGCAGTGTTCGCCGTAGGCTGCTCATCGGATATGATTGACGGGATCTCTCGGG 1201
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300

DB 1202 CTCACCTGCGCCTGTGGGACGTGTATCTGTAGAGGCAACAGCGCTGTGATGCGGATA 1261
QY 301 ThrArgIleAlaPheIysValGlnGlnIlyAspLeuThrIysThrSerArgCysGlyPro 320
DB 1262 ACAAGAAATCGCCTTTAAGGTTCCAGCAGAAAGCGCTCACGAAAGACGTCCAGGTGTGCCCCG 1321
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1322 TGGGCACGTTTTCACACCGGTTCTTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1381
QY 341 LysHISLeuAlaArgAlaSerMetLysIysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1382 AAGCATCTTAGGGCCTCTATGAGAACTAACAAAGAAAGCAGGGGACCTGCAACCCCA 1441
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
DB 1442 GCCAAACCGAGCAAGGTCGTGGGATCCAGGCTGTGGCGCTTCACGTGGCGGGAAG 1501
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1502 ACCCTCTGCAGGGGGACAGGAGCCCTCCAGGC-----CCA 1540
QY 401 IleTrpSerAlaSerProProAlaAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1541 TT-TGTCAGCTTCCCGCCAGCGGCACTCGTCTTCCACACCCTGTCTGTGGGGCT 1599
QY 421 ValArgGluAspThrTyrrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1600 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1659
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1660 GGAGGACCTCAGGGTTCCTGAGATTCTGACGTGAACCTCATGCCCGCTCCCAACG 1719
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrrAspPheArgGlnSerCysTrpVal 480
DB 1720 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGGGGCTC 1779
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
DB 1780 CGTGCCATATCCAGGAGGAGCACGCTGCCCCCTGTGGCAGGCTGAACAACCTCTCGGAG 1839
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerArgGlnGlyThrProPheArg 520
DB 1840 CGGGTGAATCGGCTTTCGTGCTGCACCACTGATTCCGACCAAGGACACCCCTTCAGA 1899
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1900 GCTAGGAGCAACAGCCGTGTGTCCCACTCAGGCGCTTGCTCTGCGGCTCCACTTG 1959
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1960 GAAAGTTCAGTTCCTCCAGGCTTC 1986
RESULT 11
ADC37378
ID ADC37378 standard, DNA; 2304 BP.
XX
AC ADC37378;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 211.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Anti-inflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Neotropic; Cardiant; Gene therapy; human; gene; ds.
OS
XX Homo sapiens.

FN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PSDB; ADC37379.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 211; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 2304 BP; 570 A; 609 C; 636 G; 489 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.52e-124 Length: 2304
Score: 2898.50 Matches: 539
Percent Similarity: 98.2% Conservative: 0
Best Local Similarity: 98.2% Mismatches: 3
Query Match: 97.2% Indels: 8
DB: 10 Gaps: 1
US-10-071-838-2 (1-549) x ADC37378 (1-2304)
QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 362 ATGACGTGTAGAGGTGCGGGCAGTTGTGTGGCAACAAGCGAGAGCATCATATATG 421
QY 21 LysTyrrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 422 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 481
QY 41 SerTyrrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 482 AGCTACAAACAACACGTGCATCATTTGGGATTGTATCATGAGACGAGCTGCTCTCTG 541
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 542 ACTGCGGGAGGCGAAGCAATTGCGCGGAGATCAGCCGAAAGACAACTGCGTACAG 601
QY 81 MetLeuGlyAspTrpGluLysTyrrLysSerSerArgLysLeuIleAspArgAlaTyrrLys 100
DB 602 ATGCTGGAGACTGGAGAAATACAAAGCAGACGAAAGCTCATAGATCGACGTACAG 661
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB 662 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACACTGAGAAATG 721
QY 121 LysLeuLysAsnProGlyArgTyrrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 722 AAGTTGAAAAACCCCGAAGATACCATGATCATGAAGAGAAAGGCAAGAGTCACTGAG 781
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160

Db 782 CACATCCAGCGCATGACCGGGACGTAAAGCGGACATTAAAGAGCATATATTCTTCAGG 841

QY 161 ASPARGTYRGLYTHRILYSGINARGGLULEULEUHSILEULEUALATYRGLUGLYYR 180

Db 842 GATCGATACGGAACCAAGCAGCGGAACTACTCACATCTCTGGCATATAGAGAGTAT 901

QY 181 ASNPROGLUVALGITYRCYSARGASPLEUSERHSILEAIAALEUPHEULEUYR 200

Db 902 AACCCGGAGGTGGGTACTGACAGGACCTGACCACATCGCCGCTTGTTCTCTCTAT 961

QY 201 LEUPROGLUGLUASPAIAPHETRPAIaleuValGlnLeuLeuAlaSerGluARGHISer 220

Db 962 CTTCTGAGAGATGCATTCTGGGCACTGGTCAGCTGCTGCCAGTAGAGGCACTCC 1021

QY 221 LEUINGLYPHEHISERPROASNGLYTHRValGlnGlyLeuGlnAspGlnGlnGlu 240

Db 1022 CTGCAGGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 1081

QY 241 HISVALVALATHRSERGLNPROLYSTHMETGLYHISGlnAspLYSLYSAspLeuCYs 260

Db 1082 CATGTGTTAGCCACTCAAAACCAAGACCATGGGCATCAGGACAGAAAGATCTATGT 1141

QY 261 GLYGLNCYSERPROLEUGLYCYSLEUIEARGILEUIEAspGLYILESERLEUGLY 280

Db 1142 GGGCAGTGTCCCGCTTAGGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 1201

QY 281 LEUTHRLAUARGLEUTRPAAPVALTYRLeuValGlnGlyGlnAlaLeuMetProIle 300

Db 1202 CTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGAACAGCGCTGATGCCGATA 1261

QY 301 THRARGILEALAPHELYVALGlnGlnLYsARGLEUTHRLYSTRHSERARGCYSGLYPro 320

Db 1262 ACAAGATCGCCTTTAAGTTACAGCAGACGCGCTCACGAAGACTCCAGGTGTGGCCCG 1321

QY 321 TRPALAARGPHECYSAANARGPHEVALASPTRTPALAAARGAspGluAspTHRValLeu 340

Db 1322 TGGGCACGTTTTCACAACGGTTCGTTGATACCTGGCCAGGATGAGACACTGTGCTC 1381

QY 341 LYSHISLEUAARGLASERMETLYSLYSELeUTHRARGLYSGInLYsAspLeuProProPro 360

Db 1382 AAGCATCTTAGGCGCTTATAGAAACTTAACAAGAAAGCAGGGGACCTGCMAACCCCA 1441

QY 361 ALALYSPROGLINGLYSERSERALASERARGPROVALPROLASERARGLYGLYLYs 380

Db 1442 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCACTGGCGGGAAG 1501

QY 381 THRLEUCYELYSGLYASPARGLNALAProProGlyProProAlaARGPheProARGPPro 400

Db 1502 ACCCTCTGCAAGGGGACAGGACGCGCTCCAGGC-----CCA 1540

QY 401 ILETTPSERALASERPROPROARGALAPROARGSERSTRHPROCYSPROGLYGLYAla 420

Db 1541 TT-TGGTCAGCTTCCCCCGCACGGGCACTGTTCTTCCACACCTGTGCTGGTGGGCT 1599

QY 421 VALARGGLUAspTHRTRYProVALGlyTHRGLNLYValProSerPROAlaLeuAlaGln 440

Db 1600 GTCCGGGAAGACACTACCTGTGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1659

QY 441 GLYGLYPROGLINGLYSERTRPARGPHEUGLINTTPASNSERMETPROARGLeuProTHR 460

Db 1660 GGAGACCTCAGGGTTCCTGAGATTCTGCAGTGAACCTCATTGCCCGCTCCCAAG 1719

QY 461 ASPLEUAspVALGLUGLYProTRPHeARGHISTYRAspPHeARGGLNSERCYSTRPVal 480

Db 1720 GACCTGAGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGCTGGTTC 1779

QY 481 ARGALALESERGLINGLYAspGlnLeuAlaProCYSTRPGLNALAGLUIHISPROAlaGlu 500

Db 1780 CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGCGAGGCTGAACACCTTGCGGAG 1839

QY 501 ARGVALARGSERALAPHEALAPROSERTHRAspSERAspGlnGlyTHRProPHeARG 520

Db 1840 CGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGAGCAGGGCACCCCTTCAGA 1899

QY 521 ALAARGAspGlnGlnProCYsAlaProTHRserGlyProCYsLeuCYsGlyLeuHISLeu 540

Db 1900 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTGTGGGCTCCACTTG 1959

QY 541 GLUSERSERGLNpHeProProGlyPhe 549

Db 1960 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1986

RESULT 12

ABZ11810

ID ABZ11810 standard; cDNA; 2084 BP.

XX AC ABZ11810;

XX AC

DT 20-JAN-2003 (first entry)

XX DE Human polynucleotide SEQ ID NO 692.

XX KW Human; genome mapping; gene therapy; food supplement; virus; fungus; cell-proliferative disorder; neurodegenerative disease; bacterial; Parkinson's disease; Alzheimer's disease; autoimmune disease; multiple sclerosis; diabetes; genetic disorder; wound; burn; infection; arthritis; cytostatic; immunomodulator; nootropic; neuroprotective; antiparkinsonian; antidiabetic; immunosuppressive; dermatological; haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide; antiarthritic; gene; ss.

XX KW

XX OS Homo sapiens.

XX PN WO200270539-A2.

XX PD 12-SEP-2002.

XX PF 05-MAR-2002; 2002WO-US005095.

XX PR 05-MAR-2001; 2001US-00799451.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F, Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M; Wehrman T, Wang J, Wang D, Drmanac RT;

XX PI

XX DR WPI; 2002-759812/82.

XX DR P-PSDB; ABP69593.

XX PT New polynucleotides comprising sequences assembled from expressed sequence tags (ESTs), useful for treating cell-proliferative, neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet or coagulation disorders.

XX PT

XX PS Claim 1; SEQ ID NO 692; 1012pp + Sequence Listing; English.

XX CC The invention relates to an isolated polynucleotide (I) comprising a nucleotide sequence selected from any of 948 sequences (ABZ1119-ABZ12066) or their mature protein coding portion, active domain coding protein or complementary sequences. The polynucleotides are useful for identifying expressed genes or for physical mapping of human genome. The encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight markers, as a food supplement, for generating antibodies, in medical imaging, screening and diagnostic assays and for treating cell-proliferative disorders (cancer), neurodegenerative diseases (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, platelet or coagulation disorders, wound, burns, incision, ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal, parasitic), arthritis, etc. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX CC

XX SQ Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.34e-122 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.4% Conservative: 1
Best Local Similarity: 96.2% Mismatches: 2
Query Match: 95.5% Indels: 18
DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x ABZ11810 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGTGTAGAGGTCCGCGGAGTTGGTGGCACAAGAGCGAGGACATCATTTATG 222
QY 21 LysTYrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 282
QY 41 SerTYrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACACAACAACGTCGATCATTTGGGATGTATCATGAGACGAGCTGCCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGGCGCAACCAATTCCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100
Db 403 ATGCTGGAGAGCTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGTCAGCGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTGAACTGAGGAAATG 522
QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGluLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGGAGCTAAGCGGACATTAAAGACATATATTCTTCAGG 588
QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db 589 GATCGATACGGAACCAAGCAGCGGGAATCTTCCATCTCTGCGCATATAGAGGATAT 648
QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTYr 200
Db 649 AACCCGGAGGTGGCTACTGACAGGACCTGAGCCATCGCCGCTTGTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCTGAGAGAGATGCATTTCTGGGCACGTGTCACGTGCTGCGCAAGTGAAGGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTGCAGGGATTTCACAGCCCAATGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAGAG 828
QY 241 HisValValAlaIleThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 949 CTCACCTGCGCCTGTGGGACGTGTATCTGTGAAGAGCGAACAAGGCTTGATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

Db 1009 ACAAGATCGCCTTTTAAGGTTCAAGCAGAAAGCCCTCACGAAGACGTCCAGGTGTGCCCC 1068
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1069 TGGCACCGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGCTCTTATGAAGAACTTAAGAAAGCAGGGGAGCTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1189 GCCAAACCCGAGCAAGGGTCGTGGCATTCAGGCTGTGCCGCTTCAAGTGGCGGAAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGAAGGGGACAGGAGGCCCTCCAGGCCACGACCCGGTTCCCGGGCCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
Db 1309 ATTTGGTCAGCTTCCCGGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGCT 1368
QY 421 ValArgGluAspThrTYrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGACACTTACCCCTGTGGGCACTCAGGGTGTGCCAGCCGGCCCTGTGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1429 GGAGACCTCAGGGTTCCTGAGATTCCTGAGTGAAGTCCATGCCCGCCCTCCCAAG 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTYrAspPheArgGlnSerCysTrpVal 480
Db 1489 GACCTGACGTAAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTG 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysSTrpglnAlaGlnHisProAlaGlu 500
Db 1549 CGTGCATATATCCAGGAGGAGCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTGGCAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CGGGTAGATCGGCTTTGCTGTCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1668
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1669 GCTAGGACGAACAGCCGTGTGCTCCACTCAGGGGCTTGCTGTGCGGCTTCCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1729 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1755
RESULT 13
ADM44328
ID ADM44328 standard; cDNA; 2084 BP.
XX
AC ADM44328;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human arginine-rich protein cDNA #692.
XX
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
XX genetic disorder.
XX
OS Homo sapiens.
XX
XX
PN US2004053250-A1.
XX
PD 18-MAR-2004.
XX
PF 21-NOV-2002; 2002US-00302172.
XX
PR 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095.

PR 20-AUG-2002; 2002US-00225251.
XX (TANG/) TANG Y T.
PA (XUEA/) XUE A.
PA (DRMA/) DRMANAC R T.
XX Tang YT, Xue A, Drmanac RT;
XX WPI; 2004-238579/22.
PT New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
PS Disclosure; SEQ ID NO 692; 51pp; English.
XX
CC The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
XX
SQ Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,34e-122 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.4% Conservative: 1
Best Local Similarity: 96.2% Mismatches: 2
Query Match: 95.5% Indels: 18
DB: 12 Gaps: 1

US-10-071-838-2 (1-549) x ADM44328 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGTGTAGAGTCCGGCGCAGTTGGTGGGCAAGAGCGAGAGCATCATTTATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTCA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 283 AGCTACACACAAACGTCGATCATTTGGGATTGTATCATGAGACGAGCTGCCTCTCTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTGCGCGGAGCGAAGCAAAATTCGGCGGAGATCAGCCGAAGCAAGTGCGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGGAGAAATACAAAGCAGACAGAAAGCTCATAGATCGAGCTACAAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACAACCTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAAGACATATTTCTTCAAG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 589 GATCGATACGAAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGATAT 648

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
Db 649 AACCCGAGGTGGCTACTGCAGGGAGCTGAGCCACATCGCCGCTTGTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCTGAGAGAGATGATTTCTGGGCATGTGTGACAGCTGTGGCCAGTGAGAGGACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGluAspGlnGlnGlu 240
Db 769 CTGCAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGACCAAGAAATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 949 CTCACCCCTGCGCCTGTGGAGCGTGTATCTGTGTAGAAGGCGAACGAGCGTTGATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATCGCCTTAAGTTACAGCAGAACGCCCTCACGAAGCGTCCAGGTGTGGCCCG 1068
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1069 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCCAAGGATGAGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1129 AAGCATCTTAGGGCCTCTATGAAGAACTTAACAAGAACGAGGGGGAAGCTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1189 GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCGGCTTCACGTGGCGGAAG 1248
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGAGACAGCAGGCCCTCCAGGCCACAGCCCGGTTCGCCGCGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
Db 1309 ATTTGTACGCTTCCCCGCCACGGGACCTCGTCTTCCACACCTGTCTGTGGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCCGGGAAGCACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1429 GGAGAACCTCAGGGTTCCTGAGATTCTGACGTGAATCCATGCCCGCTCCCAAG 1488
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1489 GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTCAGACAGAGTGTGGGTCTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1549 CGTGCCATATCCAGAGGACGACGCTGCCCTGTGGAGCTGAACACCTGCGGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1609 CGGGTAGATCGGCTTTCGCTGCACCCACGACTGATTCCGACCAAGGCAACCCCTTCAAG 1668
QY 521 AlaArgAspGluGlnProCysAlaAlaProThrSerGlyProCysLysLeuHisLysLeu 540
Db 1669 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGTCTGCGGCTTCCACTTG 1728

QY 541 GluSerSerGlnPheProGlyPhe 549
Db 1729 GAAAGTTCTCAGTTCCTCCAGGCTTC 1755

RESULT 14
ADC37388
ID ADC37388 standard; DNA; 2647 BP.

AC ADC37388;
XX 18-DEC-2003 (first entry)
XX

DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 221.

XX Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

XX Homo sapiens.

XX WO2003048202-A2.

XX 12-JUN-2003.

XX 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.

XX 05-DEC-2001; 2001US-0335829P.

XX 03-OCT-2002; 2002JP-00291302.

XX 04-OCT-2002; 2002US-0415769P.

XX (ASAH) ASAH KASEI KK.

XX Matasuda A, Muramatsu S;

XX WPI; 2003-505282/47.

XX P-PSDB; ADC37389.

XX New purified protein that activates nuclear factor kappa B (NF-kappaB),

XX useful for treating inflammation, autoimmune diseases, cancers,

XX infectious diseases, bone diseases, AIDS, neurodegenerative diseases or

XX ischemic disorders.

XX Claim 4; SEQ ID NO 221; 938bp; English.

XX The present invention relates to novel proteins and their coding

XX sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-

XX kappaB). The proteins and their coding sequences are useful for treating

XX a disease associated with NF-kappaB activation, such as inflammation,

XX autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,

XX neurodegenerative diseases, or ischaemic disorders.

XX Sequence 2647 BP; 608 A; 766 C; 750 G; 523 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 1.43e-121 Length: 2647

XX Score: 2835.00 Matches: 526

XX Percent Similarity: 95.8% Conservative: 0

XX Best Local Similarity: 95.8% Mismatches: 1

XX Query Match: 95.0% Indels: 22

XX DB: 10 Gaps: 1

US-10-071-838-2 (1-549) x ADC37388 (1-2647)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20

Db 755 ATGCAGCTGTAGAGGTGCGGGCAGTTGGTGGGCACAAGAGCGAGAGCATCATTTATG 814

QY 21 LysTYrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40

Db 815 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTAAGCCTTTTGA 874

QY 41 SerTYrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 875 AGCTACAACAACAAGCTGCATCTTTGGGGATTGTACATGAGACGAGCTGCCTCTCTG 934

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 935 ACTGCGCGGAGCGCAAGCAAAATTCGCGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 994

QY 81 MetLeuGlyAspTrpGluLysTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100
Db 995 ATGCTGGAGACTGGGAGAAATACAAAGCAGAGCAAAAGCTCATAGATCGAGCGTACAAAG 1054

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1055 GGAATGCCCAATGAACATCCGGGCGCGATGTGTCACTCTCTGAACACTGAGAAATG 1114

QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 1115 AAGTTGAAAAACCCCGGAAGATACCATCATGAAGAGAGAGGCAAGAGTCATCTGAG 1174

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAAAGCATATATCTTCAGG 1234

QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db 1235 GATCATACGGAAACCAAGCAGCGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT 1294

QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTYr 200
Db 1295 AACCCGAGGTGGCTACTGACAGGACCTGAGCCACATCGCCCTGTCTCTCTAT 1354

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 1355 CTTCTGAGGAGATGATCTTGGGCACTGTGACAGTGTGGCAGTAGAGGACATCC 1414

QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1415 CTGACAGGATTTCAACAGCCCAATGGCGGAGCGTCCAGGGGCTCCAAGACCAACAGAG 1474

QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 1475 CATGTGTAGCCACGTCAACAACCAACCATATGGGGCATCAG----- 1516

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1517 -----ATCTCTCTGGG 1528

QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 1529 CTCACCCCTGGCCTGTGGACGTATCTGTAGAAAGGCAACAGCGTTGATGCGATA 1588

QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1589 ACAAGATCGCCTTTAAGTTACAGAGAGCGCTCACAGAAAGCTCCAGGTGTGCCCCG 1648

QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1649 TGGGACAGTTTTTGGCAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC 1708

QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1709 AAGCATCTTAGGGCCTCTATGAAGAACTTAACAAGAAAGAGGGGAGCACTGCCACCCCA 1768

QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1769 GCCAAACCCGACAAAGGTGTGTGCGCATCCAGGCGCTGTGCGGCTTCACGTGCGGAAAG 1828

QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTGTGAAGGGGAGACAGGAGCGCCCTCCAGGCGCCACAGCCGCTTCCCGCGGCC 1888

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 763 CTTCCTGAGAGATGATGCTGGGCACGTGTGCACCTGTGGCCAGTGAAGGCACTCC 822
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 823 CTGCAGGGATTTCACAGCCCAATGCGGGACCGTCCAGGGCTCCAAAGACCAACAGAG 882
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 883 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 925 -----ATCTCTCTCGGG 936
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 937 CTCACCTGCGCTGTGGGACGTGTATCTGTGTAAGGCGAAACAGCGCTTGATGCCGATA 996
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 997 ACAAGAATCGCCTTTAAGGTTCAAGCAGAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1057 TGGGCACGTTTTTGCAACGGTTGCTGATACCTGGGCCAGGGATGAAGACACTGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1117 AAGCATCTTAGGGCCTCTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCAACCCCA 1176
QY 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1177 GCCAAACCCGAGCAAGGCTGTGCGCATCCAGGCTGTGCCGCTTCACTGCGGGAG 1236
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1237 ACCCTCTGCMAAGGGGACAGGCAGGCCCTCCAGGCCACAGCCCGCTTCCCGGGCCC 1296
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1297 ATTTGTCAGCTTCCCGGCCACGGGCACCTCGTTCTTCCACACCTGTCTGTGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1357 GTCCGGGAAGACACCTAACCTGTGGCACCTCAGGGTGTGCCAGCCGGCCCTGCTCAG 1416
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1417 GGAGGACCTCAGGGTTCCTGGAGATTCTGTGCATGGAACTCCATGCCCGCCTCCCAACG 1476
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1477 GACCTGACGTAGAGGGCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1536
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 1537 CGTGCCATATCCCAAGAGAACAGCTGGCCCCCTGTGCAGGCTGAACACCTGTCCGAG 1596
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1597 CGGGTGAGATCGGCTTTCGTGCACCCAGCACCTGATTCGACCAAGGCACCCTTCAGA 1656
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1657 GCTAGGGACGAACAGCCGTGTCTCCCACTCAGGGCCTTGCCCTGTGCGGCTCCACTTG 1716
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1717 GAAAGTTCTCAGTTCCTCCAGGCTTC 1743

Search completed: April 5, 2006, 14:49:21
Job time : 1222 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 5, 2006, 14:42:06 ; Search time 6702 Seconds
(without alignments)
4656.379 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWWAQEREDIM.....TSGPCLGHLHSSQPPPGF 549

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/abs/ABSSWEB_spool/US10071838/runat_05042006_141913_5146/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs805p
-USER=US10071838@CGC_1_1_7415@runat_05042006_141913_5146 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2983	100.0	1964	6	AX775945	AX775945 Sequence
2	2983	100.0	1964	8	AF540953	AF540953 Homo sapi
3	2975	99.7	1964	6	AX775947	AX775947 Sequence

4	2971	99.6	2090	8	BC071680	BC071680 Homo sapi
5	2963	99.3	2072	6	AX086847	AX086847 Sequence
6	2945	98.7	2166	8	BC075809	BC075809 Homo sapi
7	2906.5	97.4	2304	6	AX775943	AX775943 Sequence
8	2898.5	97.2	2304	6	AX775941	AX775941 Sequence
9	2850	95.5	2084	6	AR578516	AR578516 Sequence
10	2849	95.5	2575	8	BC033670	BC033670 Homo sapi
11	2835	95.0	2647	6	AX775951	AX775951 Sequence
12	2827	94.8	2072	6	AR578515	AR578515 Sequence
13	2827	94.8	2647	6	AX775949	AX775949 Sequence
14	2827	94.8	2647	6	AX833552	AX833552 Sequence
15	2827	94.8	2647	8	AK095385	AK095385 Homo sapi
16	2732	91.6	2111	8	BC058890	BC058890 Homo sapi
17	2626.5	88.0	1924	8	BC078140	BC078140 Homo sapi
18	2595	87.0	3389	8	AK122833	AK122833 Homo sapi
19	2191.5	73.5	7878	6	AX332512	AX332512 Sequence
20	2191.5	73.5	7878	8	HSTR210	HSTR210
21	2188.5	73.4	7971	8	AY143550	AY143550 Homo sapi
22	2184.5	73.2	8491	8	HSMB07865	HSMB07865
23	2133.5	71.5	8201	6	I76205	I76205 Sequence
24	2133.5	71.5	8201	6	AX774956	AX774956 Sequence
25	2133.5	71.5	8201	8	HSTR213	HSTR213
26	1522	51.0	134506	14	AC146482	AC146482 Homo sapi
27	1522	51.0	180876	8	AC021317	AC021317 Homo sapi
28	1522	51.0	181517	8	AC126327	AC126327 Homo sapi
29	1521	51.0	193646	8	AC131056	AC131056 Homo sapi
30	1518	50.9	146243	14	AC130293	AC130293 Homo sapi
31	1514	50.8	37757	8	AC139091	AC139091 Homo sapi
32	1505	50.5	173053	14	AC129893	AC129893 Homo sapi
33	1501	50.3	140152	8	AC027821	AC027821 Homo sapi
34	1501	50.3	140152	8	AC027821	AC027821 Homo sapi
35	1500.5	50.3	158698	14	AC036181	AC036181 Homo sapi
36	1500.5	50.3	192881	8	AC067923	AC067923 Homo sapi
37	1500	50.3	206040	14	AG132814	AG132814 Homo sapi
38	1497.5	50.2	185449	14	AC139361	AC139361 Homo sapi
39	1494.5	50.1	42073	14	AC144840	AC144840 Homo sapi
40	1493.5	50.1	41265	14	AC139176	AC139176 Homo sapi
41	1493.5	50.1	89447	14	AC146857	AC146857 Homo sapi
42	1487.5	49.9	146243	14	AC130293	AC130293 Homo sapi
43	1469	49.2	123925	8	AC003976	AC003976 Homo sapi
44	1465	49.1	172790	14	AC023133	AC023133 Homo sapi
45	1448.5	48.6	163628	14	AC139367	AC139367 Homo sapi

ALIGNMENTS

RESULT 1	AX775945	1964 bp	mRNA	linear	PAT 14-JUL-2003
LOCUS	AX775945				
DEFINITION	Sequence 215 from Patent WO03048202.				
ACCESSION	AX775945				
VERSION	AX775945.1	GI:32693663			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Homidae; Homo.				
REFERENCE	1				
AUTHORS	Matsuda, A. and Muramatsu, S.				
TITLE	NF-kB activating gene				
JOURNAL	Patent: WO 03048202-A 215 12-JUN-2003;				
	Asahi Kasei Kabushiki Kaisha (JP)				
FEATURES	Location/Qualifiers				
source	1..1964				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	1..1650				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAE11672.1"				
	/db_xref="GI:32693664"				

/translation="MDVVEVAGSWWAQEREDIMKYEKHBAGLPEDKGPFPRESVNN
NVDHLGI VHETELPLTAREAKQIRREISRSKSWVDMLGDMEKYKSSRKLIDRAYKGM
PMNIRGPMWSVLINIEEMKLVGRYQIMKEKGRSSEHORIDRDVSGTLRKHIFER
DRYGTOKORELHILILAYBEYNPEVGYCRDLSHIALFLYLPEBDAFMALVQLASER
HSLQGFHS PNGGTVQGLQDOQEHVAVTSQKTMGHODKDLGGCSPGLCLIRLIDG
ISLGLTRLMDVYLVVEGEOLMPTRIAFKVQOKRLTKTSRCGPMARFCNRFVDTMAR
DEDTVLKHLSRASMKKLTRKQGDLPAPAKBEQSSASRPVPASRGKTLCKGDROAPG
PPARFPRIPIWSASPPPARSRSTPBGCAVREDTYPVGTQGVSPALAOGGPOGSRFL
QWNSMPRLPTDL DVEGPMFRHYDFRGWVRALISQEDQLPCMQAEHPAERVRSAFLA
PSTSDQGTFFRARDEQPCAPTSGLCLHLESSQFPFGF"

ORIGIN

Alignment Scores:

Pred. No.:	9.43e-95	Length:	1964
Score:	2983.00	Matches:	549
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-071-838-2 (1-549) x AX775945 (1-1964)

QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
Db	1	ATGACGTGCTAGAGGTGCGGGGCAAGTTGGTGGGCAACAAGCGAGGACATCATATATG	60
QY	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
Db	61	AAATACGAAGAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCGA	120
QY	41	SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu	60
Db	121	AGCTACACACACACACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCCTTG	180
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
Db	181	ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAAGCCGAAAGAGCAAGTGGTGAT	240
QY	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
Db	241	ATGCTGGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAA	300
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet	120
Db	301	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTCGAACATTGAGGAATG	360
QY	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
Db	361	AAATTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGGAAGGCAAGGCTCATCTGAG	420
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
Db	421	CACATCCAGCGCATCGACCGGAGCTTAAGCGGACATTAAAGAACATATATTCTTCAGG	480
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
Db	481	GATCGATACGGAACCAAGCAGCGGAACTACTCCATCCTCCTGGCATATGAGGAGTAT	540
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr	200
Db	541	AAACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTCTCTCTAT	600
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	601	CTTCCTGAGGAGGATGCTTCTGGGCACTGTGTGCACTGTGCGCCAGTGAAGGCACTCC	660
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	661	CTGCAGGGAATTCACAGCCCAAAATGGCGGACCTCCAGGGGCTCCAAAGACCAAGAG	720
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	721	CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGACAAAGAAATCTATGT	780

QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	781	GGGCAGTGTTCCCCGTTAGGCTGCTCATCCGGATATTGATTGACGGGATCTCTCGGG	840
QY	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle	300
Db	841	CTCACCTCGGCTGTGGACGTGTATCTGTAGAAAGGCGAACAGCGCTTGATGCCGATA	900
QY	301	ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro	320
Db	901	ACAAGATCGCTTTTAAGGTTCAAGCAGAACCGCTCACGAAGACGTCCAGGTGTGCCCG	960
QY	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	961	TGGGCACGTTTTCGAACCGGTTGCTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC	1020
QY	341	LysHisLeuArgAlaSerMetLysLeuLeuThrArgLysGlnGlyAspLeuProProPro	360
Db	1021	AAGCATCTTAGGGCTCTATGAGAACTAAACAAGAACGAGGGGACCTGCCACCCCCA	1080
QY	361	AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys	380
Db	1081	GCCAAACCCGAGCAAGGCTGTCCGCTCCAGGCTGTGCTGCCCTTCACTGCGGGAG	1140
QY	381	ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1141	ACCCTTGCAAGGGGACAGGACAGGCCCTCCAGGCCACAGCCCGGTTCCCGCGGCC	1200
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla	420
Db	1201	ATTGGTCACTTCCCGCCACGGGCACTCGTTCTTCCACACCTGTCTGTGGGCT	1260
QY	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1261	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGTTGTGCCAAGCCCGCTGGCTCAG	1320
QY	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	1321	GAGGACCTCAGGTTCTCTGGAATTCCTGCACTGGAATCCATGCCCGCTCCCAACG	1380
QY	461	AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
Db	1381	GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC	1440
QY	481	ArgAlaIleSerGlnLysAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu	500
Db	1441	CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTGGAG	1500
QY	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
Db	1501	CGGGTAGATCGGCTTTCGCTGACACCACTGATTCGACACAGGACCCCTTCAG	1560
QY	521	AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu	540
Db	1561	GCTAGGACGAAACAGCCGTGTGCTCCACCTCAGGCGCTTGTGCGGCTCCACTTG	1620
QY	541	GluSerSerGlnPheProProGlyPhe	549
Db	1621	GAAGTTCTCAGTTCTCTCCAGGCTTC	1647

RESULT 2
AF540953 1964 bp mRNA linear PRI 02-OCT-2002
LOCUS
DEFINITION Homo sapiens Rab GTPase-activating protein PRC17 (PRC17) mRNA,
complete cds.
ACCESSION AF540953
VERSION AF540953.1 GI:23452664
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 1964)
AUTHORS Pei,L., Peng,Y., Ling,X.B., van Eynhoven,W.G., Nguyen,K.C.Q., Rubin,M., Hoey,T., Powers,S. and Li,J.
TITLE PRC17, a novel oncogene encoding a Rab GTPase-activating protein, is amplified in prostate cancer
JOURNAL Cancer Res. (2002) In press.
REFERENCE 2 (bases 1 to 1964)
AUTHORS Li,J.L., Peng,Y. and Powers,S.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2002) Genomics, Tularik, Inc, 266 E Pulaski Road, Suite 1, Greenlawn, NY 11740, USA

FEATURES
source 1. 1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
1. 1964
/gene="PRC17"
1. 1650
/gene="PRC17"
/codon_start=1
/product="Rab GTPase-activating protein PRC17"
/protein_id="AAN3117.1"
/db_xref="GI:23452665"
/translation="MDVVEVAGSWMVAQEREDIMKYEKGHAGLPEDKGPFRSYNN
NVDHLGI VHETELPPLTAREAKQIRREISRSKMWDMGDWEKYSRKLIDRAYKGM
PMNIRGPMWSVLINIEMKLNKNGRYQIMKEGKRSSEHIQRIDRVSGTLAKHIFER
DRYGTQRELIHLILAYEEYNEVGYCRDLSHIALFLIYLBEDAFWALVQLASER
HSLGFHSPNGVTQGLQDQGEHVAVTSQPKTMGHQDKDLGGCSPICGLRILIDG
ISLGLTLRLMDVYLVEGEQALMPITRIAFKVQKRLTTSRCGPWARFCNRFVDTWAR
DEDTVLKHLRASMKGLTRKQGLDLPAPAKBEGSSASRPVPASRGKTLCKEDROAPG
PPARFPRPIWSASPPRAPRSSTPCPGGAVREDTYPVGTQGVSPALAQGGPSWRFL
QWNSMPRLPTDLDEGPWFRRHDFRQSCWRAISQEDQLAPCWQAEHPAERVRSAPAA
PSTDSDQGTFRARDEQPCAPTSRPLCGHLLESSQFPBGF"

ORIGIN

Alignment Scores:
Pred. No.: 9, 43e-95 Length: 1964
Score: 2983.00 Matches: 549
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-071-838-2 (1-549) x AF540953 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 1 ATGGACGTGGTAGAGCTCGCGGGCAGTTGGTGGGCACAAGAGCGAGACATCATTTATG 60
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 61 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGGCTTAAGCCTTTTCGA 120
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 121 AGCTACCAACAACGTCGATCATTTGGGGAATTGTACATGACGAGCGAGCTGCCTCTTG 180
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 181 ACTGCGCGGAGCGAAGCAAAATTCGGCGGAGATCAGCCGAAGAGCAAGTGCGTGAT 240
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 241 ATGCTGGAGACTCGGAGAATAACAAGACAGCAAGAACTCATAGATCGAGCGTACAG 300
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAAACATTGAGGAATG 360

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 361 AAGTTGAAAACCCCGAAGATACCATCATGAAAGAGAAAGGCAAGAGTCATCTGAG 420
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCCAGCGCATCGACCGGACGCTAAGCGGACATTAAGAAACATATATCTTCAGG 480
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 481 GATCGATACCGAACCAAGCAGCGGAACTACTCCACATCTCTGCGCATATGAGGAGTAT 540
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
Db 541 AACCGGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCCTTGTCTCTCTTAT 600
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 601 CTTCCTGAGGAGATGCATTTGGGCACTGTGACAGCTGTGGCCAGTGAGAGGCACTCC 660
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 661 CTGCAGGATTTACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 720
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 721 CATGTGTTAGCCACGTACACAAACCAACCATGGGGCATCAGACAAGAAAGATCTATGT 780
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCTGGG 840
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnGlnAlaLeuMetProIle 300
Db 841 CTCACCTCGCGCTGTGGACGTGTATCTGTAGAAGCGCAACAGCGTGTATGCCGATA 900
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 901 ACAAGAATCGCCTTTAAGTTCAAGCAGAACGCCCTCAGCAAGACGTCCAGGTGTGCGCCG 960
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 961 TGGGCACGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGATGAGGACATGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1021 AAGCATCTTAGGGCTCTATAGAAATACTAACAGAAAGCAGGGGCACTGCCACCCCCA 1080
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1081 GCCAAACCCGACCAAGGTCGTGGCATCCAGGCTGTGGGCTTCACGTGCGGGAAG 1140
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1141 ACCCTCTGCAAGGGGGAACAGCAGGCCCTCCAGGCCACAGCCGGTTCCCGGGGCC 1200
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1201 ATTTGGTCAGTTCCCGGCCACGGGCACTCGTTCTTCCACACCTGTCTGTGGGGCT 1260
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1261 GTCCGGGAAGACACTTACCCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGCTCAG 1320
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1321 GGAGACCTTCAGGGTTCCTGGAGATTCCTGAGTGAATTCATGCCCGCTCCCAACG 1380
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1381 GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTCAGACAGAGCTGTGGGTC 1440
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500

Db 1441 CGTGCATATCCAGAGGACCAAGCTGCCCCCTGCTGGCAGGCTGACACACCTGCGAG 1500

QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520

Db 1501 CGGGTGAGATCGGCTTCCTGCTGCACACCAAGCACTGATTCGCAACCAAGGACACCCCTTCAGA 1560

QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCySGlyLeuHisLeu 540

Db 1561 GCTAGGAGACGAACAGCCGTGTGCTCCACACCTCAGGGCCTTGCTCTGCGGCTCCACTTG 1620

QY 541 GluSerSerGlnPheProGlyPhe 549

Db 1621 GAAAGTTCTCAGTTCCTCCAGGCTTC 1647

RESULT 3

AX775947 1964 bp mRNA linear PAT 14-JUL-2003

LOCUS AX775947

DEFINITION Sequence 217 from Patent WO03048202.

ACCESSION AX775947

VERSION AX775947.1 GI:32693665

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1

AUTHORS Matsuda, A. and Muramatsu, S.

TITLE NF-kB activating gene

JOURNAL Patent: WO 03048202-A 217 12-JUN-2003;

Asahi Kasei Kabushiki Kaisha (JJP)

FEATURES

source

1. 1964

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1. 1650

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE11673.1"

/db_xref="GI:32693666"

/translation="MDVBEVAGSWMWAQEREDIIIMKYEKGRAGLPEDKGPKEPERSYNN

NVDHLGIVHETELPPLTAREAKQIRREISRKSKWVMDLGDWEKYKSSRKLIDRAYKGM

PMNIRGPMWSVLLNIEEMKIKNPGYQIMKEKGKRSSEHQRIDRDVSGTLRKHIFER

DRYGTQKQRELIHLILAYEENPEVGYCRDLSHIALFLYLPEEDAFWALVQLASER

HSIQGFHSPNGSTVQGLDQDQEHVAVTSQPKTMGHQDKDLCGQCSPLGCLIRLLIDG

ISLGLTLRLMDVYLVVEEQALMPLTRIAFKVQKRLTKTSRCGPWARFCNRFYDTWAR

DEDTVLKHILASMKKLTRKQGLDPPAKPEQSSASRPVASRGKTLCKGDRQAPG

PPARFPRPIWSASPPRAPRSSTPCGAVREDTYPVGTGVPSPALAQGGPOGSWRL

QWNSMRLPTDLDVEGPFRIHYDRQSCWVAISQEDQLAPCWQAEHPAERVSAFAA

PSTSDSQTFPRARDEQPCAPTSBCLCGLHLESSQFPQGF"

CDS

1. 1650

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE11673.1"

/db_xref="GI:32693666"

/translation="MDVBEVAGSWMWAQEREDIIIMKYEKGRAGLPEDKGPKEPERSYNN

NVDHLGIVHETELPPLTAREAKQIRREISRKSKWVMDLGDWEKYKSSRKLIDRAYKGM

PMNIRGPMWSVLLNIEEMKIKNPGYQIMKEKGKRSSEHQRIDRDVSGTLRKHIFER

DRYGTQKQRELIHLILAYEENPEVGYCRDLSHIALFLYLPEEDAFWALVQLASER

HSIQGFHSPNGSTVQGLDQDQEHVAVTSQPKTMGHQDKDLCGQCSPLGCLIRLLIDG

ISLGLTLRLMDVYLVVEEQALMPLTRIAFKVQKRLTKTSRCGPWARFCNRFYDTWAR

DEDTVLKHILASMKKLTRKQGLDPPAKPEQSSASRPVASRGKTLCKGDRQAPG

PPARFPRPIWSASPPRAPRSSTPCGAVREDTYPVGTGVPSPALAQGGPOGSWRL

QWNSMRLPTDLDVEGPFRIHYDRQSCWVAISQEDQLAPCWQAEHPAERVSAFAA

PSTSDSQTFPRARDEQPCAPTSBCLCGLHLESSQFPQGF"

ORIGIN

Alignment Scores:

Pred. No.: 1.78e-94 Length: 1964

Score: 2975.00 Matches: 548

Percent Similarity: 99.8% Conservative: 0

Best Local Similarity: 99.8% Mismatches: 1

Query Match: 99.7% Indels: 0

DB: 6 Gaps: 0

US-10-071-838-2 (1-549) x AX775947 (1-1964)

QY 1 MetAspValValGluValAlaGlySerTrpAlaGlnGluArgGluAspIleIleMet 20

Db 1 ATGGAAGTGTAGAGGTCCGGGCAAGTGTGGGCAACAAGAGCGAGAGCATCATTTATG 60

QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40

Db 61 AAATACGAAAGGACACCGAGCTGGGCTGCCAGACGACAAAGGGCCTTAAGCCTTTTGA 120

QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60

Db 121 AGTACAACAACAAGCTGATCATTTGGGGATTGTACTAGACGAGCTGCCCTCTCTG 180

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

Db 181 ACTGCCGGGAGCGGAGCAAGCAAAATTCGCGGGAGATCAGCCGAAAGACAACTGGGTGAT 240

QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100

Db 241 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAGAAAGCTCATAGATCGAGCGTACAAG 300

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120

Db 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTACATCTCTGAACATTGAGGAAATG 360

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140

Db 361 AAGTTGAAAACCCCGGAAGATACAGATCATGAAGGAGAAGGCAAGAGTCACTGAG 420

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160

Db 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGAAAGCATATATCTTCAGG 480

QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluIleTyr 180

Db 481 GATCGATACGGAACCAAGCAGCGGGAATTACTCCACATCTCTCTGCAATATGAGAGTAT 540

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200

Db 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTTCTCTCTAT 600

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

Db 601 CTTCTGAGGAGATGATCTTGCGCACTGTGACAGCTGTGCGCAGTGAAGGCACTCC 660

QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240

Db 661 CTGCAAGGATTTCAACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 720

QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260

Db 721 CATGTGTAGCCACGCTCACAAACCAAGACCATGGGGCATCAGGACAAAGATCTATGT 780

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280

Db 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840

QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300

Db 841 CTCACCCCTGGCCTGTGGACGTGTATCTGTAGAAAGGGAACAGCGTTGATGCGATA 900

QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

Db 901 ACAAGAATGCGCTTTAAGGTTCAAGAGCGCCTCAGCAAGACGTCCAGGTGTGCCCG 960

QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340

Db 961 TGGGACGTTTTTGGCAACCGGTCTGTGATACCTGGGCAAGGATGAGACACTGTGCTC 1020

QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360

Db 1021 AAGCATCTTAGGGCTCTATGAAGAACTTAACAAGAAAGCAGGGGAGCCTGCACACCCCA 1080

QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380

Db 1081 GCCAAACCGGAGCAAGGGTGTGTCGATCCAGGCTGTGCGGCTTACAGTGGCGGGAAG 1140

QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400

Db 1141 ACCCTCTGCAAGGGGAGACAGGAGGCCCTCCAGGCCCCACAGCCGGTTCCCGGGCCC 1200

QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420

Db 1201 ATTTGTCAGCTTCCCCGACAGGACACCTGTTCTTCCACACACCTGTCTGTGGGCT 1260

Qy 421 ValArgGluAspThrTyrProValGlyPheGlnGlyValProSerProAlaLeuAlaGln 440

Db 1261 GTCCGGGAAGACACTACCTGTGGGACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320

Qy 441 GlyGlyProGlnGlySerTyrPargPheLeuGlnTyrPasnSerMetProArgLeuProThr 460

Db 1321 GGAGGACCTCAGGGTTCCTGGAGATTCTTGCAGTGGAACCTCCATGCCCCCTCCCAACG 1380

Qy 461 AspLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480

Db 1381 GACCTGACCTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1440

Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrPglAlaGluHisProAlaGlu 500

Db 1441 CGTGCCATATCCAGAGACAGACAGCTGGCCCCCTGTGACAGGCTGAACACCCCTGCGAG 1500

Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520

Db 1501 CGGTGAGATCGGCTTTCGCTGCACCCAGACTGATTCGACAGAGGACACCCCTTCAGA 1560

Qy 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540

Db 1561 GCTAGGAGCAGACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTG 1620

Qy 541 GluSerSerGlnPheProProGlyPhe 549

Db 1621 GAAAGTTCTCAGTTCCTCCAGGCTTC 1647

RESULT 4

BC071680 2090 bp mRNA linear PRI 08-MAR-2005

LOCUS Homo sapiens TBC1 domain family, member 3, mRNA (cDNA clone

DEFINITION MGC:87891 IMAGE:5165385), complete cds.

ACCESSION BC071680

VERSION BC071680.1 GI:47939522

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2090)

AUTHORS Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Bueltow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Ustin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SM, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butcherfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.

CONSRM Mammalian Gene Collection Program Team

TITLE Generation and initial analysis of more than 15,000 full-length

JOURNAL human and mouse cDNA sequences

PUBMED Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

REFERENCE 12477932

AUTHORS 2 (bases 1 to 2090)

TITLE Director MGC Project.

JOURNAL Direct Submission

Submitted (01-JUN-2004) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-bhg.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAC Plate: 166 Row: g Column: 13

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.

FEATURES

source

1. .2090

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:87891 IMAGE:5165385"

/issue_type="Brain, adult medulla"

/clone_lib="NIH MGC_119"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .2090

/gene="TBC1D3"

/note="synonyms: PR17, TRE17"

/db_xref="GeneID:84218"

/db_xref="MIM:607741"

101. .1750

/gene="TBC1D3"

/codon_start=1

/product="TBC1 domain family, member 3"

/protein_id="AAH71680.1"

/db_xref="GI:47939523"

/db_xref="GeneID:84218"

/db_xref="MIM:607741"

CDS

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

ORIGIN

Alignment Scores:

Pred. No.: 2.6e-94 Length: 2090

Score: 2971.00 Matches: 546

Percent Similarity: 99.8% Conservative: 2

Best Local Similarity: 99.5% Mismatches: 1

Query Match: 99.6% Indels: 0

DB: 8 Gaps: 0

US-10-071-838-2 (1-549) x BC071680 (1-2090)

Qy 1 MetaAspValAlaGluValAlaGlySerTyrTrpAlaGlnGluArgGluAspIleIleMet 20

Db 101 ATGACGCTGTAGAGTCCGGGCACTTGTGGGACACAGACGAGAGACATCATTTATG 160

Qy 21 LysTyrGluGlyGlyHisArgAlaGlyLeuProGluAspGlyGlyProLysProPheArg 40

Db 161 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 220

Qy 41 SerTyrAsnAsnAsnValaAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60

Db 221 AGCTACACAAACAAGTCAATTTGGGATTGTATCATGAGACGAGCTGCTCTCTG 280

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 281 ACTGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGACAAAGTGGTGAT 340
QY 81 MetLeuGlyAspTrpGlnLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 341 ATGCTGGAGACTGGGAGAAATACAAAAGACAGAAAGCTCATAGATCGAGCGTACAAG 400
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluIleMet 120
Db 401 GGAATGCCCATGAACATCCGGGGGCCGATGTGTCAGTCTCTGAACATTGAGGAATG 460
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 461 AAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGGCAAGGTCATCTGAG 520
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 521 CACATCCAGCGCATGACCGGGACATTAAGCGGACATTAAAGAACATATGTTCTTCAGG 580
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGlyTyr 180
Db 581 GATCGATACGGAACCAAGCAGCGGAACCTACTCCACATCCTCCTGCAATATGAGAGTAT 640
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 641 AACCCGAGGTGGGCTACTGCAGGACCTGAGCCATCGCCGCTGTCTCTCTCAT 700
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 701 CTTCTGAGAGAGATGCATTTCTGGGCACTGGTGCAAGTGTGCGCAAGAGGCACTCC 760
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 761 CTGCAGGGATTTCACAGCCCAAAATGGCGGACGTCACAGGGGCTCCAAGACCAAGAGAG 820
QY 241 HisValValAlaIleHisSerGlnProLysIleHisGlnAspLysLysAspLeuCys 260
Db 821 CATGTGTAGCCACGTCACAACCCAAAGACCATGGGGCATCAGGACAGAAAGATCTATGT 880
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 881 GGGCAGTGTCCCGTTAAGCTGCTGCCTCATCGGATATTGATTGACGGGATCTCTCGGG 940
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 941 CTCACCCTCGCGCTGTGGGACGTGTATCTGTGAAGAAGCGAACAAGCGTTGATGCCGATA 1000
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1001 ACAAGATCGCCTTAAAGGTTCAAGCAAGACCGCTCAAGAAAGCTCCAGGTGTGGCCCG 1060
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1061 TGGGCACGTTTTCGAACCGGTTGTTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1120
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1121 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGACAGGGGACCTGCCACCCCCA 1180
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1181 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCCGCTCAAGTGGCGGAAG 1240
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1241 ACCCTCTGCAAGGGGACAGGAGGCCCTCTCAAGGCCCAAGCCCGGTTCCCGCGGCC 1300
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1301 ATTGGTCAAGCTTCCCGCCACGGGCACTGTTCTTCACACACCTGTCTGTGGGGCT 1360
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

Db 1361 GTCCGGAGAGACACCTACCCCTGTGGGCACTCAGGGTGTGCCAGCCCGCCCTGGCTCAG 1420
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1421 GGAGGACCTCAGGGTTCCTGGAGATTCCTGCAGTGAAGTCCATGCCCGCCCTCCCAAG 1480
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1481 GACCTGGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1540
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 1541 CGTGCCATATCCAGAGAGACCAAGCTGGCCCTGCTGGCAGGCTGAACACCCCTGGAG 1600
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1601 CGGTGAGATCGGCTTTCGTGTCACCCAGCACTGATTCGACCAAGGCACCCCTTCAGA 1660
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1661 GCTAGGAGACAAAGCAGAGTGTGCTCCACTCAGGGCCTTGCTCTGCGGCTCCACTTG 1720
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1721 GAAAGTTCAGTTCCTCCAGGCTTC 1747

RESULT 5
AX086847 LOCUS AX086847 2072 bp DNA linear PAT 09-MAR-2001
DEFINITION Sequence 799 from Patent WO0112659.
ACCESSION AX086847
VERSION AX086847.1 GI:13276047
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Wiemann,S.
TITLE Human dna sequences
JOURNAL Patent: WO 0112659-A 799 22-FEB-2001;
German Human Genome Project (DB)
FEATURES
source Location/Qualifiers
1..2072
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.86e-94 Length: 2072
Score: 2963.00 Matches: 545
Percent Similarity: 99.6% Conservative: 2
Best Local Similarity: 99.3% Mismatches: 2
Query Match: 99.3% Indels: 0
Gaps: 0
US-10-071-838-2 (1-549) x AX086847 (1-2072)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 99 ATGACGTGTGTAGAGGTGCGCGGCAAGTTGTGTGGCACAGAAGCGAGAGCATCATTTATG 158
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 159 AAATACGAAGAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA 218
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 219 AGCTACAACAACAGTCGATCATTTGGGGATTGTACATGAGACGAGAGCTGCCTCTCTG 278

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerIleTrpValAsp 80
Db 279 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGGTGAT 338
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 339 ATGCTGGAGACTGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 398
QY 101 GlyMetPrometAsnIleArgGlyPrometTrpSerValLeuLeuAsnIleGluMet 120
Db 399 GGAATGCCCATGAAACATCCGGGGCCCGATGTGTCATCTCTGAAACACTGAGGAATG 458
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 459 AAGTTGAAAAACCCCGGAAGATACCATCATGAAAGGAAGGGCAAGAGTCACTCTGAG 518
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 519 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAGCATATATTCTTCAGG 578
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 579 GATCGATACGGAACCAAGCAGCGGAACTACTCATCTCCTGGCATATGAGAGATAC 638
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleAlaPheLeuLeuTyr 200
Db 639 AACCCGGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 698
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 699 CTTCTGAGAGGATGATTTCTGGCAGCTGTGTCAGCTGTGCTGAGAGAGCACTCC 758
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 759 CTGACGGGATTTCAGACCCCAATAGCGGACGCTCCAGGGGCTCCAGAACCAACAGAGAG 818
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 819 CATGTGTAGCCACGTCACAACCCAGACCAATGGGGCATCAGACAGAAAGATCTATGT 878
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 879 GGGCAGTGTCCCGCTTAGGCTGCCTCATCCGATATTGATTGACGGGATCTCTCCGG 938
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 939 CTCACCCCTCGCCTGTGGGACGTGTATCTGTAGAAAGGCGAACAGCGCTGATGCCGATA 998
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 999 ACAAGATCGCCTTTAAGTTACAGCAGAGGCCCTCACGAAGACGTCAGTGTGCCCG 1058
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1059 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1118
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1119 AAGCATCTTAGGGCTCTATGAGAACAATAAGAAAGAGGGGACCTGCCACCCCCA 1178
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1179 GCCAAACCCGAGCAAGGGTCTCGGCATCCAGGCCCTGTGCCGCTTCACGTGGCGGAAG 1238
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1239 ACCCTCTGCAAGGGGACAGGCAGGCCCTCTCCAGGCCACACGACCGGTTCCCGCGGCC 1298
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1299 ATTTGGTCAGCTTCCCGCCACGCGGACACTCGTTCTTCACACACCTGTCTGTGGGGCT 1358
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

Db 1359 GTCCGGGAAGACACTTAACCTCTGTGGCAGCTAGGGGTGTGCCAGCCCGCCCTGCTCAG 1418
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1419 GAGAGACCTCAGGGTTCTTGAGATTCTCTGCACTGGAACCTCATGCCCCGCTCCCAACG 1478
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1479 GACCTGACGATAGAGGCCCTTGGTTCGCCCATATGATTTCAGACAGAGCTGTGGTTC 1538
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1539 CGTGCCATATCCAGAGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTCGCGAG 1598
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1599 CGGGTGAGATCGGCTTTCCTGCTGCACCCAGCACTGATTCCGACCAAGGCAACCCCTCAGA 1658
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1659 GCTAGGAGCAAGACAGAGTGTGCTCCCACTCAGGGCCTTGCTGCGGCTCCACTTG 1718
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1719 GAAAGTTCTCAGTTCCTCCACAGGCTTC 1745

RESULT 6
BC075809 2166 bp mRNA linear PRI 17-MAY-2005
LOCUS BC075809
DEFINITION Homo sapiens TBC1 domain family, member 3B, mRNA (cDNA clone
MGC:87892 IMAGE:5170064), complete cds.
ACCESSION BC075809
VERSION BC075809.1 GI:49902676
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 2166)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRM Mammalian Gene Collection Program Team
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2166)
AUTHORS NIH MGC Project
CONSRM Direct Submission
TITLE Submitted (06-JUL-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
JOURNAL NIH-MGC Project URL: http://mgc.nci.nih.gov
REMARK Contact: MGC help desk
COMMENT Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@pax1.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAK Place: 166 Row: h Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.
Location/Qualifiers

FEATURES

source

1. .2166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87892 IMAGE:5170064"
/tissue_type="Brain, adult medulla"
/clone_1b="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2166
/gene="TBC1D3B"
/note="synonym: PRC17"
/db_xref="GeneID:414059"
147. .1796
/gene="TBC1D3B"
/codon_start=1
/product="TBC1 domain family, member 3B"
/protein_id="AAH75809.1"
/db_xref="GI:49902677"
/db_xref="GeneID:414059"
/translation="MDVVEVAGSWAQEREDIMKYEKGRAGLPEDKGPKPRSYNN
NVDHLGIVHETELPLTAREAKQIRREISRSKSWVDMLGMEKYKSRKLIBRAYKGM
PMNIRGPMWSVLNIEEMKLNPGRYQIMKEGKRSEHIQIDRISGTLRRHMF
DRYGTQRELIHLILAYEENPEVQCRDLSHIALFLYLPEDAFWALVQLASER
HSLQGFHSPNGSTVQGLQDQDEHVAVTSQSTMGHDKDLGQCSPLGCLIRLIDR
ISLGLTLRLMDVYLVEGEALMPTRIAFKVQQRLLTKTSRCGPWFARFCNRFVDTWAR
DDDTVLKHLRASMKLTRKQGLDPPAKPEQSSASRPVPSRGKTLCKGDRQAPG
PPARFPRPIWSASPPPARPRSTPCPGAVREDTYPVGTQVPSPALAQGGPQGSWRFL
QWNSMPRLPTDLVDEGPMFRHYDFEQSCWVRASIQEDQLCWMQAEHPAERVSAAFAA
PGTSDQGTPTFRARDEQQCAPTSGPCLCGLHLESSQFPPEF"

CDS

1. .2166
/gene="TBC1D3B"
/note="synonym: PRC17"
/db_xref="GeneID:414059"
147. .1796
/gene="TBC1D3B"
/codon_start=1
/product="TBC1 domain family, member 3B"
/protein_id="AAH75809.1"
/db_xref="GI:49902677"
/db_xref="GeneID:414059"
/translation="MDVVEVAGSWAQEREDIMKYEKGRAGLPEDKGPKPRSYNN
NVDHLGIVHETELPLTAREAKQIRREISRSKSWVDMLGMEKYKSRKLIBRAYKGM
PMNIRGPMWSVLNIEEMKLNPGRYQIMKEGKRSEHIQIDRISGTLRRHMF
DRYGTQRELIHLILAYEENPEVQCRDLSHIALFLYLPEDAFWALVQLASER
HSLQGFHSPNGSTVQGLQDQDEHVAVTSQSTMGHDKDLGQCSPLGCLIRLIDR
ISLGLTLRLMDVYLVEGEALMPTRIAFKVQQRLLTKTSRCGPWFARFCNRFVDTWAR
DDDTVLKHLRASMKLTRKQGLDPPAKPEQSSASRPVPSRGKTLCKGDRQAPG
PPARFPRPIWSASPPPARPRSTPCPGAVREDTYPVGTQVPSPALAQGGPQGSWRFL
QWNSMPRLPTDLVDEGPMFRHYDFEQSCWVRASIQEDQLCWMQAEHPAERVSAAFAA
PGTSDQGTPTFRARDEQQCAPTSGPCLCGLHLESSQFPPEF"

ORIGIN

Alignment Scores:

Pred. No.: 2.11e-93 Length: 2166
Score: 2945.00 Matches: 541
Percent Similarity: 99.3% Conservative: 4
Best Local Similarity: 98.5% Mismatches: 4
Query Match: 98.7% Indels: 0
DB: 8 Gaps: 0

US-10-071-838-2 (1-549) x BC075809 (1-2166)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 147 ATGACGCTGTAGAGTCGCGGGTAGTTGTGGGCAACAAGACGAGGACATCATATATG 206
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 207 AAATAACGAAAGGACACCGAGCTGGGCTGCCAGAGACAGGCGCTTAAGCCTTTTCA 266
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 267 AGCTACAAACAACACGTCATCTTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 326
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 327 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGACAGTGCGTGAT 386

QY 81 MetLeuGluAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 387 ATGCTGGAGACTGGGAGAATAACAAGACAGAAAGCTCATAGATGAGCGTACAAG 446
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
DB 447 GGAATGCCCATGAACATCCGGGGCCCATGTGTCACTCTCTGAACATTGAGGAATG 506
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
DB 507 AAGTGAATAACCCCGAAGATACGATCATGAAGAGAAAGGCAAGAGTCACTGAG 566
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 567 CACATCCAGCGCATCGACCGGACATTAAGCGGACATTAAGAGCATATGTTCTTCAGG 626
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 627 GATCGATACGGAACAAGACAGCGGGACTACTCCACATCCTCTGGCATATGAGAGTAT 686
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
DB 687 AACCCGAGAGTGGGCTACTGCAAGGACCTGAGCCACATCGCGCCTGTCTCTCTAT 746
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 747 CTTCCTGAGGAGATGCATTCTGGGCACTGTGCAAGCTGCTGCCAAGTAGAGGCACTCC 806
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
DB 807 CTGACGGAATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACGAG 866
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 867 CATGTGTAGCCACGTCACAATCCAGACCATGGGCGATCAGCAAGAAAGATTAATGT 926
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 927 GGGCAGTGTTCCTCGTTAGCTGCTCATCCGATATTGATGACCGATCTCTCGGG 986
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
DB 987 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAGGCGCAACAGCGTTGATGCCGATA 1046
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
DB 1047 ACAAGATCGCCTTTAAGTTACAGCAAGCGCCTCACGAAGACGTCCAGGTGTGCCCA 1106
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1107 TGGCACGTTTTCACAACCGGTCGTGTGATACCTGGGCCAGGATGATGACACTGTGCTC 1166
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1167 AAGCATCTTAGGGCCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1226
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
DB 1227 GCCAAACCCGACAAAGGTCGTCCGATCCAGGCTGTGCCGCTTACAGTGGCGGAAG 1286
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1287 ACCCTCTGAAAGGGGACAGGACAGGCCCTCCAGGCCACCAAGCCGGTCCCGCGGCC 1346
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1347 ATTGGTCAAGTTCCTCCGACGGGACCTGTTCTTCCACACCCTGTCTGTGGGCT 1406
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1407 GTCCGGAAGACACCTACCTGTGGCACTCAGGGTGTGCCACCGCGCCTGTGCTCAG 1466

Qy 441 G|YGIProG|NG|YSerTrpArgPheLeuG|N|TrpAsnSerMetProArgLeuProThr 460
Db 1467 GGAGGACCTCAGGGTTCCTGAGATTCTGCAGTGAACCTCATGCCCCGCTCCCAAG 1526
Qy 461 AspleuAspValG|U|YProTrpPheArgHisTyrAspPheArgG|N|SerCysTrpVal 480
Db 1527 GACCTGAGCGTAGAGGGGCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGTC 1586
Qy 481 ArgAlaIleSerG|N|G|U|AspG|N|LeuAlaProCysTrpG|N|AlaG|U|HisProAlaG|U| 500
Db 1587 CGTGCCATATCCACAGAGGACCAAGCTGCCCCCTGTGGCAGCTGAACACCCCTGGCAG 1646
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspG|N|YThrProPheArg 520
Db 1647 CGGGTGAGATCGGCTTTCGCTGCACCCGGCACTGATTCGACCAAGGCAACCCCTTCAGA 1706
Qy 521 AlaArgAspG|U|G|N|ProCysAlaProThrSerG|Y|ProCysLeuCysG|Y|LeuHisLeu 540
Db 1707 GCTAGGAGCAACAGACAGTGTGCTCCCACTCAGGGGCTTGCTGCGGCTCCACTTG 1766
Qy 541 GluSerSerG|N|PheProProG|Y|Phe 549
Db 1767 GAAAGTCTCAGTTCCTCCAGGCTTC 1793
RESULT 7
AX775943 2304 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775943 Sequence 213 from Patent WO03048202.
DEFINITION AX775943
ACCESSION AX775943
VERSION AX775943.1 GI:32693661
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 213 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. 2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11671.1"
/db_xref="GI:32693662"
/translation="MDVVEVAGSWAQEREDIMKYEKGRAGLPBDKGPKPRSYNN
NVDHLGIVHETELPPLTAREAKQIRREISRKSMVDMLEWKEKSSRKLIDRAYKM
PMNIRGPMWSVLNTEEMKLNKPRYQIMKEGKRSSHHIQRIDRVSGLRKHIFER
DRYGTQRELEHILAYEENPEVGYCRDLSHIALFLYLPEEDAFWALVQLASER
HSLQGFSNNGTVQGLQDQEHVATISQPTWGHQDKDLGGCSPGLCLIRILIDG
ISLGLRLMDVYLVEGEQALMPITRIAFYQQKRLTKTSRCGPWAFRCNRFVDTMAR
DEDTVLKLRLASMKLTKRQGLPPPAKPEQSSASRPVPASRGKTLCKGDRQAPPG
PFGQLPRHGHVLPHPVLVGLSGKPTLMALRVCARPMLREDLRVPGDSCSGTPCPA
SORTWT"
ORIGIN
Alignment Scores:
Pred. No.: 4.76e-92 Length: 2304
Score: 2906.50 Matches: 540
Percent Similarity: 98.4% Conservative: 0
Best Local Similarity: 98.4% Mismatches: 2
Query Match: 97.4% Indels: 8
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AX775943 (1-2304)
Qy 1 MetAspValValG|U|ValAlaG|Y|SerTrpTrpAlaG|N|G|U|ArgG|U|AspIleIleMet 20

Db 362 ATGAGCTGTGTAGAGCTCGCGGGCAGTTGTGGGCACAAAGCCGAGAGCATCATATAG 421
Qy 21 LysTyrG|U|LysG|Y|HisArgAlaG|Y|LeuProG|U|AspLysG|Y|ProLysProPheArg 40
Db 422 AAATACGAAGAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTCA 481
Qy 41 SerTyrAsnAsnAsnValAspHisLeuG|Y|IleValHisG|U|ThrG|U|LeuProProLeu 60
Db 482 AGCTACAAACAACAGTCGATCATTTGGGATTGTACATGAGACGAGCTGCCCTCTG 541
Qy 61 ThrAlaArgG|U|AlaLysG|N|IleArgArgG|U|IleSerArgLysSerLysTrpValAsp 80
Db 542 ACTGCGCGGAGCGGAGCAATTCCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 601
Qy 81 MetLeuG|Y|AspTrpG|U|LysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 602 ATGCTGGGAGACTGGGAGAAATACAAAGACAGCAAGAACTCATGATCGAGCGTACAG 661
Qy 101 GlyMetProMetAsnIleArgG|Y|ProMetTrpSerValLeuLeuAsnIleG|U|Met 120
Db 662 GGAATGCCATGAACATCCGGGCGCGATGTGTGATCTCTCTGAACACTGAGGAAATG 721
Qy 121 LysLeuLysAsnProG|Y|ArgTyrG|N|IleMetLysG|U|LysG|Y|LysArgSerSerG|U| 140
Db 722 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 781
Qy 141 HisIleG|N|ArgIleAspArgAspValSerG|Y|ThrLeuArgLysHisIlePhePheArg 160
Db 782 CACATCCAGCGCATCGACCGGGACGTAAAGCGGACATTAAGGAAGCATATATCTTCAGG 841
Qy 161 AspArgTyrG|Y|ThrLysG|N|ArgG|U|LeuLeuHisIleLeuLeuAlaTyrG|U|G|U|Tyr 180
Db 842 GATCGATACGGAACCAAGCAGCCGGAATCTACTCCACATCTCTCGCATATGAGAGTAT 901
Qy 181 AsnProG|U|ValG|Y|TyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 902 AACCCGAGGTGGCTACTGACGGAACCTGAGCCACATCGCCGCTGTCTCTCTAT 961
Qy 201 LeuProG|U|G|U|AspAlaPheTrpAlaLeuValG|N|LeuLeuAlaSerG|U|ArgHisSer 220
Db 962 CTTCTGAGAGAGATGATCTGTGGCACTGTGTGCACTGTGGCCAGTGAAGGCACTCC 1021
Qy 221 LeuG|N|G|Y|PheHisSerProAsnG|Y|G|Y|ThrValG|N|G|Y|LeuG|N|AspG|N|G|N|G|U| 240
Db 1022 CTGACGGGATTTACAGACCCAAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 1081
Qy 241 HisValValAlaThrSerG|N|ProLysThrMetG|Y|HisG|N|AspLysLysAspLeuCys 260
Db 1082 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 1141
Qy 261 G|Y|G|N|CysSerProLeuG|Y|CysLeuIleArgIleLeuIleAspG|Y|IleSerLeuG|Y| 280
Db 1142 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 1201
Qy 281 LeuThrLeuArgLeuTrpAspValTyrLeuValG|U|G|Y|G|U|AlaLeuMetProIle 300
Db 1202 CTCACCTCGCGCTGTGGACGTGTATCTGTGAAGGCGAAGAGCGCTGTGATGCCGATA 1261
Qy 301 ThrArgIleAlaPheLysValG|N|G|N|LysArgLeuThrLysThrSerArgCysG|Y|Pro 320
Db 1262 ACAGAAATCGCCTTTAAGTTACAGACAGAGCGCTCAAGAAAGCTCCAGGTGTGGCCCG 1321
Qy 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspG|U|AspThrValLeu 340
Db 1322 TGGGCACTTTTGTCAACCGGTTCTGTGATACCTGGGCCAGGAGTGAAGACACTGTGCTC 1381
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysG|N|G|Y|AspLeuProProPro 360
Db 1382 AAGCATCTTAGGGCTCTATGAAGAACTAACAAGAAAGCAGGGGAGCTGCCACCCCA 1441
Qy 361 AlaLysProG|U|G|N|G|Y|SerSerAlaSerArgProValProAlaSerArgG|Y|G|Y|Lys 380

Db	1442	GCCAAACCCGAGCAAGGGTCGTCCGCATCCAGGCCCTGTGCCGCTTCACGTCGGCGGAAG	1501
Qy	381	ThrLeuCySlySGlyAspArgGlnAlaProProGlyPProProAlaArgPheProArgPro	400
Db	1502	ACCTCTGCAGAGGGGACAGGACGGCCCTCCAGGC-----CCA	1540
Qy	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCySProGlyGlyAla	420
Db	1541	TT-TGGTCAGCTTCCCCGCCACGGGCACTCGTCTTCCACACCCTGTCTCGTGGGGCT	1599
Qy	421	ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1600	GTCCGGGAGACACCTACCTGTGGCACTCAGGGTGTGCCAGCCCGCCCTGGCTCAG	1659
Qy	441	GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	1660	GGAGGACCTCAGGGTCTCTGAGATTCTGCAGTGAATCATGCCCCCTCCCAACG	1719
Qy	461	AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
Db	1720	GACCTGACGTAGAGGGCCCTGTGTCGCCATTATGATTTCAGACAGAGCTGTGGGTC	1779
Qy	481	ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu	500
Db	1780	CGTGCCATATCCCAAGAGACCAAGCTGCCCCCTGTGGCAGGCTGAACACCCCTGCGAG	1839
Qy	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
Db	1840	CGGGTGAAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCGCACCCCTCAGA	1899
Qy	521	AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCySglyLeuHisLeu	540
Db	1900	GCTAGGAGACAAACGCCGTGTGCTCCCACTCAGGGCCTTGCTTGGCGGCTCCACTTG	1959
Qy	541	GluSerSerGlnPheProGlyPhe	549
Db	1960	GAAAGTTCAGTTCTCCCTCCAGGCTTC	1986
RESULT 8			
AX775941		2304 bp	mRNA
LOCUS	AX775941	Sequence 211 from Patent WO03048202.	linear
DEFINITION	AX775941		PAT 14-JUL-2003
ACCESSION	AX775941		
VERSION	AX775941.1	GI:32693659	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 Matsuda, A. and Muramatsu, S.		
TITLE	NF-kB activating gene		
JOURNAL	Patent: WO 03048202-A 211 12-JUN-2003;		
	Asahi Kasei Kabushiki Kaisha (JP)		
FEATURES			
source	1..2304		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	362..1732		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAE11670.1"		
	/db_xref="GI:32693660"		
CDS			
	translation="MDVVEVAGSWMAOEREDIMKYEKGHRAGLPEDKGKPKFRSYNM		
	NVHLGI VHETELPPLTAREAKOIRBISRSKSNVMDLGWEKYSSRKLDRA YKGM		
	PMNIRGPMWSVLNTEEMKLNPRGYQIMKEGKRSEHIORIDRDVSGTIRKGIIFR		
	DRYGTQRELHLHLAYEENPEVGYCRDLSHIALFLYLPEDAFWALVQLLASER		
	HSLGFFHSPNGGTGVGLQDQGEHVAVATSPKTMGHODKDKDLGQCSPLGLIRLIDG		
	ISLGLTLRLMDVYLVESQALMPITRIAFKVQOKRLTTSRCGPWAPFCNRFVDTWAR		
	DEDYTLKHLRASMKLTKRGDIOPRAPKEGSSASRPVPASRGKTLCKGDRQAPRG		
	PFGQLPRHGHVLVLPVVLVGLSGKPTLWALRVCPARPWLREDLRVPGDSCSGTPCPA		

ORIGIN	SQRTWT"
Alignment Scores:	
Pred. No.:	8.98e-92
Score:	2898.50
Percent Similarity:	98.2%
Best Local Similarity:	98.2%
Query Match:	97.2%
DB:	6
	Gaps: 1

US-10-071-838-2 (1-549) x AX775941 (1-2304)

Qy	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
Db	362	ATGACGTGTGTAGAGTCCGGGCAAGTTGGTGGGCAAGAGCGAGACATCATTAAG	421
Qy	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
Db	422	AAATACGAAAAAGGACACCGAGCTGGGCTGCAGAGGACAAAGGGCCTTAAGCTTTTCCA	481
Qy	41	SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu	60
Db	482	AGCTACAAACAACAGCTGCATGATTTGGGATTTGTACATGAGACGAGCTGCTCTTG	541
Qy	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
Db	542	ACTCGCGGGAGCGGAAAGCAATTCCGCGGGAGATCACCCGAAAGCAAGTCGGTGAT	601
Qy	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
Db	602	ATGCTGGAGACTGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCTACAAAG	661
Qy	101	GlyMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGluMet	120
Db	662	GGAATGCCATGAACATCCGGGCGCCGATGTGTCACTCTCTGAACTGAGGAAATG	721
Qy	121	LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
Db	722	AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGGAGAAGGCAAGAGTCTCTGAG	781
Qy	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
Db	782	CACATCCAGCGCATGCACCGGACGTAAAGCGGACATTAAGAAACATATATCTTCAGG	841
Qy	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
Db	842	GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTCTGCATATGAGAGTAT	901
Qy	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr	200
Db	902	AACCCGGAGGTGGCTACTGCAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT	961
Qy	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
Db	962	CTTCTGAGAGAGATGATTCGTGGCACTGTGCAGCTGCTGGCCAGTAGAGGCACTCC	1021
Qy	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	1022	CTGCAGGAGATTTCACAGCCCAATGGCGGAGACCGTCCAGGGGCTCCAAGACCAAGAG	1081
Qy	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
Db	1082	CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAAGACAAGAAAGATCTATGT	1141
Qy	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	1142	GGGAGAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG	1201
Qy	281	LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle	300
Db	1202	CTCACCTGCGCCTGTGGAGCGTGTATCTGTAGAAAGCGAAACAGCGTTGATCCGATA	1261

QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1262 ACAAGAATCGCCTTAAAGTTACAGACAGAGCGCTCACGAAGACGTCAGGTGGCCCG 1321
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1322 TGGGACGCTTTTGCAACCGGTTCGTTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1381
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlnLysAspLeuProPro 360
Db 1382 AAGCATCTTAGGGCTCTATGAGAACTAACAGAAAGCAGGGGACCTGCACACCCCA 1441
QY 361 AlaLysProGlnGlnGlnLysSerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1442 GCCAAACCCGAGCAAGGGTGTGTCGATCCAGGCTGTGCCGCTTCACTGGCGGGAAG 1501
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1502 ACCCTCTGCAAGGGGACAGGCAGGCCCCCTCCAGGC-----CCA 1540
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1541 TT-TGGTACGCTTCCCCGCCACGGGCACCTGTTCTTCCACACCTGTCTGTGGGGCT 1599
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1600 GTCCGGGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGCCCTGGGCTCAG 1659
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 1660 GGAGGACCTCAGGGTTCCTGAGATTCTGCAGTGAACTCCATGCCCGCCCTCCCAAG 1719
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 1720 GACCTGACGTAGAGGGCCCTGTGTTCCGCATTATGATTTCAGACAGAGCTGTGGGTG 1779
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 1780 CGTGCCATATCCCAAGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCCGAG 1839
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1840 CGGGTGAATCGGCTTTCGCTGCACCCACCACTGATTCCGACCAAGGGCAACCCCTTCAGA 1899
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 1900 GCTAGGGACGAACAGCCGTGTGCTCCCACTCAGGGCTTGTGCTCTGCGGCTCCACTTG 1959
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 1960 GAAAGTTCTCAGTTCCTCCAGGCTTC 1986
RESULT 9
AR578516
LOCUS AR578516 2084 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 692 from patent US 6783969.
ACCESSION AR578516
VERSION AR578516.1 GI:56581312
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2084)
AUTHORS Tang, Y. T., Goodrich, R. W., Aeundi, V. and Drmanac, R. T.
TITLE Cathepsin V-like polypeptides
JOURNAL Patent: US 6783969-A 692 31-AUG-2004;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
source location/Qualifiers
1..2084
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.: 3,83e-90 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.4% Conservative: 1
Best Local Similarity: 96.2% Mismatches: 2
Query Match: 95.5% Indels: 18
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AR578516 (1-2084)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 163 ATGACGTGTAGAGGTCCGGGCGACGTTGTGGGCACAAAGCGAGACATCATTTATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AAATACGAAAGGAGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 282
QY 41 SerTyrAsnAsnValAspHisLeuGlyIleValHisGlnThrGluLeuProProLeu 60
Db 283 AGCTACACACACACGTCGATCATTTGGGGATTGTACATGACGAGAGCTGCCTCCTTG 342
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 343 ACTCGCGGAGGCGGAAGCAAAATTCGCGGGAGATCAGCCGAAAGCAAGTGGGTGAT 402
QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGACTGGGAGAAATACAAAGACAGACAAAGCTCATAGATCGAGCTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACACTGAGGAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 523 AAGTTG----- 528
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAGGAACATATATTCTTCAGG 588
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 589 GATCGATACGAACCAAGCAGCGGAACTACTCCACATCTCTCGCATATGAGGATAT 648
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
Db 649 AACCCGAGGTGGGCTACTGCAGGGAACCTGAGCCACATCGCCGCTTGTCTCTCTAT 708
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCTGAGAGAGATGATTCATTCGGGCACTGTGACGTGCTGGCCAGTGAAGGCACTCC 768
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 769 CTGACGGGATTTTCAACGCCCAATGGCGGAGCCGTCAGGGGGCTCCAAAGACCAAGAG 828
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTAGCCACGTCAACCAACCAAGACCATGGGGCATCAAGCAAGAAAGATCTATGT 888
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 889 GGGCAGGTTCCTCCGTTAGGCTGCTCATCCGATATGATTGACGGGATCTCTCGGG 948
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
Db 949 CTACACCTCGCGCTGTGGAGCGTATCTGTAGAAAGCGAAACAGCGTGTGATGCCGATA 1008
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGAATCGCCTTAAAGTTACAGACAGAGCGCTCAGAAAGAGCTCAGGTGTGGCCCG 1068

QY 321 TrpAlaArgPheCyAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
|||||
Db 1069 TGGGCACGTTTTTGCAACCGGTTGTTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1128
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
|||||
Db 1129 AAGCATCTTAGGGCTCTATGAGAACTACAGAAGACAGGGGGACCTGCAACCCCA 1188
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
|||||
Db 1189 GCCAAACCCGAGCAAGGGTCGTCGGCATCCAGGCTGTGCCGCTTCACTGGCGGGAAG 1248
QY 381 ThrLeuCyLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
|||||
Db 1249 ACCCTCTGCAAGGGGACAGGACAGGCCCTTCAGGCCACACCGCGGTTCCCGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCySProGlyGlyAla 420
|||||
Db 1309 ATTTGGTCAGCTTCCCGCCACGCGACCTCGTCTTCCACACCTGTCTGTTGGGCT 1368
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
|||||
Db 1369 GTCCGGGAAGACACCTAACCTGTGGCAGCTCAGGGTGTGCCAGCCGGCCCTGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
|||||
Db 1429 GGAGGACCTCAGGGTCTCTGGAGATTCTCTGAGTGAACCTCCATGCCCGCTCCCAACG 1488
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
|||||
Db 1489 GACCTGACGTAAGGGGCCCTTGTTCCGCATTATGATTTCAGACAGAGCTGTGGGTC 1548
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
|||||
Db 1549 CGTGCATATCCACAGAGGACAGCTGGCCCTGCTGTCAGAGCTGAACACCCCTGCGAG 1608
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
|||||
Db 1609 CGGGTGAGATCGGCTTCCGTGCACACCACTGATTCCGACAGGGCAACCCCTTCAGA 1668
QY 521 AlaArgAspGluGlnProCyAlaProThrSerGlyProCyLysGlyLeuHisLeu 540
|||||
Db 1669 GCTAGGAGCAGACCGGTGTGCTCCCACTCAGGGCCTTGCTGCGGCTCCACTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
|||||
Db 1729 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1755

RESULT 10
BC033670 2575 bp mRNA linear PRI 09-MAY-2005
LOCUS BC033670 2575 bp mRNA linear PRI 09-MAY-2005
DEFINITION Homo sapiens TBC1 domain family, member 3C, mRNA (cDNA clone
MGC:44903 IMAGE:5167394), complete cds.
ACCESSION BC033670
VERSION BC033670.1 GI:45708750
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2575)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Cabavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Utsin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boufard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2575)
PUBMED
JOURNAL
TITLE
CONSRM
AUTHORS
CONSRM
TITLE
JOURNAL
REMARK
COMMENT

NIH MGC Project
Direct Submission
Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Boufard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stanciripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgion,C., Vogt,J.L., Walker,M.A., Wecherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAX Plate: 68 Row: 1 Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 48949891.
location/Qualifiers

FEATURES
source
1. .2575
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:44903 IMAGE:5167394"
/tissue_type="Brain, adult medulla"
/clone_1ib="NIH_MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2575
/gene="TBC1D3C"
/note="synonym: MGC44903"
/db_xref="GeneID:414060"
78. .1208
/gene="TBC1D3C"
/codon_start=1
/product="TBC1 domain family member 3C"
/protein_id="AAH33670.1"
/db_xref="GI:45708751"
/db_xref="GeneID:414060"
/translation="MDVVEVAGSWAQERDIIIMKEYEGHRAGLPEDKGPFRSYNN
NYDHLGIHBTLPPLTARBAKQIRREISRSKWMDMGWEKYSRKLIDRAYKGM
PNINIRGPMWSVLNIEEMKLNPRYQIMKEGRSSEHIQIRIDRISGTLRKMFER
DRYTKQRELIHLLAYEYNEPVGCRDLSHIALFLYLPEDAFVALVQLASER
HSLQGFHSPNGTVQDQDQEHVATISQKTMGHQDKDLGQCSPLGLIRILIDG
VRHRETLAGPSLPCALLLPQGGLAHSQPTGSGSGSPKDTQAKPSAQEGSSQGNW
GSGPASWADWARTLERAQSLKPWASPLOEPHPSNECP"

ORIGIN

Alignment Scores:

Pred. No.:	5.08e-90	Length:	2575
Score:	2849.00	Matches:	543
Percent Similarity:	77.4%	Conservative:	2
Best Local Similarity:	77.1%	Mismatches:	4
Query Match:	95.5%	Indels:	155
DB:	8	Gaps:	1

US-10-071-838-2 (1-549) x BC033670 (1-2575)

QY	1	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
DB	78	ATGAGCGTGTAGAGGTCCGGGTAGTGTGGGCAACAAGCGAGGACATCATTTATG	137
QY	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
DB	138	AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA	197
QY	41	SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu	60
DB	198	AGCTACACACACACGTCGATCATTTGGGGATTGTACATAGACGGAGCTCCCTCTG	257
QY	61	ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp	80
DB	258	ACTCGCGGAGGCGCAAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT	317
QY	81	MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys	100
DB	318	ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAAAAGCTCATAGATCGACGTACAAG	377
QY	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet	120
DB	378	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCAGTCCCTGAACATTGAGGAATG	437
QY	121	LysLeuLysAsnProGluYArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu	140
DB	438	AAGTGAAAAACCCCGAAGATACCAAGATCATGAAGGAAGGCAAGGTCATCTTGAG	497
QY	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg	160
DB	498	CACATCCAGCGCATCGACCGGGACATAGCGGGACATTAAAGGACATATGTTCTTCAG	557
QY	161	AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr	180
DB	558	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAT	617
QY	181	AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr	200
DB	618	AACCCGAGGTGGCTACTGCAGGGACCTGAGCCACATCGCGCTTGTCTCTCTAT	677
QY	201	LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer	220
DB	678	CTTCCTGAGGAGATGCATTCTGGGCACCTGTGCAAGCTGTGGCCAGTGAGGGCACTCC	737
QY	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlu	240
DB	738	CTGCAGGGATTTCACAGCCCAATATGGCGGACCGTCCAGGGGCTCCAAAGACCAACAGAG	797
QY	241	HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys	260
DB	798	CATGTGTAGCCACGTCACAATCCAAAGCCATGGGCATCAGGACAAAGAAAGATCTATGT	857
QY	261	GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAsp-Gly-----	276
DB	858	GGGCAGTGTTCGCCGTTAGGCTGCCTCATCCGGATATTGATTGACGGGGTAAGAGGCAT	917
QY	276	-----	276
DB	918	AGGAGACCTGGCTCAGGGAACCTTCCCTTGCCTGCAGTGCCCTGTCCCAAGCCGGG	977
QY	276	-----	276

DB	978	GGTCTGGCTACTCCACGCCACAGAGGCTCAGCGGGTCCCAAAAGACACACAAGCA	1037
QY	276	-----	276
DB	1038	AAACCTCTGCCAAGAGGGGTCAATCCAGGGCAATGTGGGGCTCAGGCCAGCCTCA	1097
QY	276	-----	276
DB	1098	TGGGACAGCTGGGCCAGGACCCGACTTGAGAGGGCTCAGGGAAGCCTCAAGCCCTGGGCA	1157
QY	276	-----	276
DB	1158	AGCCCTCTCTCCAAGAGCCACATCCCACTCAATGATGCCCCCATGAGAGCTTCA	1217
QY	276	-----	276
DB	1218	AGACCTGTGTGACCCAGCGTCTGAGGGCTCAGCCGACCTCATGGGGAAGTCACTG	1277
QY	276	-----	276
DB	1278	ACTCTGAGACTGAAGCCCAAGCGTGCAGCTCGAGCCACCAAGCCCAAGCTGAAGGA	1337
QY	277	-----IleSerLeuGlyLeuThrLeuArgLeuTr	286
DB	1338	CCAGGTTCTTTCACACCTGCTGTCCCAAGATCTCTCGGGCTCACCTGCGCTGTG	1397
QY	286	PAspValTyrLeuValGluGlyGlnAlaLeuMetProIleThrArgIleAlaPheLys	306
DB	1398	GGAAGTATCTGTGTAGAAAGCGCAAGCGCTTGATGCCGATTAACAAGATCGCTTTAA	1457
QY	306	SValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTrpAlaArgPheCysAs	326
DB	1458	GGTTCAGCAGAGCGCCTCAGCAAGAGCTCAGGTGTGGCCCGTGGGACGTTTTCGA	1517
QY	326	NArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLysHisLeuArgAlaSe	346
DB	1518	CCGGTTCGTTGATACCTGGGCAAGGATGAGACACTGTGTCTCAAGCATCTTAGGGCTTC	1577
QY	346	rMetLysLysLeuThrArgLysGlnGlyAspLeuProProProAlaLysProGluGlnI	366
DB	1578	TATGAAGAACTTAACAAGAAAGCAGGGGACCTGCCACCCCAAGCAAGCAAGG	1637
QY	366	ySerSerAlaSerArgProValProAlaSerArgGlyLysThrLeuCybLysGlyAs	386
DB	1638	GTCGTGGGATCCAGGCCCTGTGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGA	1697
QY	386	pArgGlnAlaProProGlyProProAlaArgPheProArgProIleTrpSerAlaSerPr	406
DB	1698	CAGGAGGCCCCCTCAGGCCACCAAGCCGGTCCCGCGGCCATTGTGTCAGCTTCCT	1757
QY	406	oProArgAlaProArgSerSerThrProCysProGlyGlyAlaValArgGluAspThrTy	426
DB	1758	GCCACGGGCACTCGTCTTCCACACCTGTCTGTGGGGGTGTCCGGGAAGACACTTA	1817
QY	426	rProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyProGlnGlySe	446
DB	1818	CCCTGTGGGCACTCAGGGGTGTGCCACCCCGGCTGTGCTCAGGAGGACCTCAGGGTTC	1877
QY	446	rTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGluI	466
DB	1878	CTGAGATTCTTGCAAGTGAATCTCATGCCCCGCTCCAAAGGACCTGAGCGTAGAGGG	1937
QY	466	yProTrpPheArgHisTyrAspPheArgGlnSerCysTrpValArgAlaIleSerGlnI	486
DB	1938	CCCTGTGTTCCGCATATATGATTTCAGACAGAGCTGTGGGTCCGTGCATATCCAGGA	1997
QY	486	uAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGluArgValArgSerAlaPh	506
DB	1998	GGACCAAGTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAGCGGGTGAGATCGGCTTT	2057
QY	506	eAlaAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGluGlnPr	526
DB	2058	CACGCAACCAAGCACTGATTCGACCAAGGGCAACCCCTTTCAGAGCTAGGGAAGAAAGCA	2117

QY 526 oCySaLaProThrSerGlyProCySbLeuCySgLyLeuHisbLeuGluSerSerGlnPhePr 546
Db 2118 GTGTGCTCCACCTTCAGGGCCCTTGCTCTGCGGCCCTCCACTTGAAAGTTCTCAGTTCCC 2177
QY 546 oProGlyPhe 549
Db 2178 TCCAGGCTTC 2187
RESULT 11
AX775951 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775951
DEFINITION Sequence 221 from Patent WO03048202.
ACCESSION AX775951
VERSION AX775951.1 GI:32693669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 221 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
Source Location/Qualifiers
1..2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11675.1"
/db_xref="GI:32693670"
/translation="MDVVEVAGSWWAQEREDIMKYEKGHRAGLPEDKGPKPFRSYNN
NVDHLGIVHETELPPLTAREAKQTRREISRSKSWVMDLGDWEKYKSSRKLDAYKGM
PMNIRGPMWSVLNTEEMKLNPNRYQIMKEKGKRSSEHQRIDRVSGLTRKHIFER
DRYGTQREBLHLILAYEENPEVGYCRDLSHIALFLYLPEBDAFWALVQLASER
HSLGFHSPNGTVQGLQDOOEHVATSOPTMGHQSGLGLRLMDVYLVEGSOALM
PTIRIAFKYQQRKLTKTSTRCCPMARFCNRFVDTWARDDETVLKHILRASMKLTRKQGD
LPPPAKEGGSSASRPVPASRGKTLCKGDRAAPGPPARFPRIWSASPPARSSST
PCPGAVREDTYPVGTQGVSPALAQGGPQGSWRFLQWNSMPRLPTDLDVEGPFRRHY
DFRQSCWVARAISQEDQLAPCWQAEHPAERVRSFAFAPSTDSDQTPFRARDEQCAPT
SGPCLCGHLHSSQFPFGF"
ORIGIN
Alignment Scores:
Pred. No.: 1.58e-89 Length: 2647
Score: 2835.00 Matches: 526
Percent Similarity: 95.8% Conservative: 0
Best Local Similarity: 95.8% Mismatches: 1
Query Match: 95.0% Indels: 22
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AX775951 (1-2647)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 755 ATGACGCTGCTAGAGTCGCGGGCAGTTGGTGGGCACAAGAGCGAGGACATCATATTATG 814
QY 21 LysTYrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 815 AAATACGAAGAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGGCTTAAGCCTTTTCGA 874
QY 41 SerTYrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 875 AGCTACAACCAACACGTCGATCATTTGGGGATTGTACATGAGACGGAAGCTGCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
Db 935 ACTGCGGAGGAGCGAAGCAAAATTGCGCGGAGATCAGCCGAAAGACAAAGTGCGGTGAT 994

QY 81 MetLeuGlyAspTrpGluLysTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100
Db 995 ATGCTGGAGACTGGGAGAAATACAAAGCAGACGAAGAAGCTCATAGTCGAGCGTACAAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1055 GGAATGCCCATGAAACATCCGGGGCCCGATGTGTCACTCTCTGAAACACTGAGGAATG 1114
QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 1115 AAGTTGAAAACCCCGGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTTCATCTGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGGAACATATATTCTTCAGG 1234
QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTYrGluGluTYr 180
Db 1235 GATCGATACGAACCAAGCAGCGGAACCTACTCCACATCCTCTGCGCATATGAGAGTAT 1294
QY 181 AsnProGluValGlyTYrCySArgAspLeuSerHisIleAlaLeuPheLeuLeuTYr 200
Db 1295 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 1355 CTTCCTGAGAGGATGATCTTGCGCACTGTGTCAGCTGCTGCGCACTGAGAGGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1415 CTGCAGGATTTCAACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAACAGAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCyS 260
Db 1475 CATGTGTAGCCACGTCACAACCCAGACCATGGGCGATCAG----- 1516
QY 261 GlyGlnCySerProLeuGlyCySLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1517 -----ATCTCTCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrpAspValTYrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 1529 CTCACCCCTGGCTGTGGACGTGATCTGTGTAAGAGCGCAACAGCGCTTGATGCCGATA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCySglyPro 320
Db 1589 ACAAGATCGCTTTAAGTTTACGAGAAAGCGCCTCACGAAGACGTCAGGTGTGCCCG 1648
QY 321 TrpAlaArgPheCySAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1649 TGGGCACGTTTTTGCAACCGGTCGTTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1709 AAGCATCTTAGGCCCTCTATGAGAACTAACAAGAAAGCAGGGGACCTGCCACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1769 GCCAAACCCGAGCAAGGGTCGTCCGCATCCAGGCCCTGTGCCGCTTCACGTGCGGGAAG 1828
QY 381 ThrLeuCySlyGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGCAAGGGGAGACAGGAGGCCCTCCAGGCCACCAAGCCCGGTTCCCGCGGCC 1888
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCySProGlyGlyAla 420
Db 1889 ATTGTGTCAGCTTCCCGCCACGGGCACTGTTCTTCCACACCTGTCTGTGGGGCT 1948
QY 421 ValArgGluAspThrTYrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCCGGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 2008

Qy	441	G G G G P r o G I n G L y S e r T r P a r g P h e L e u G I n T r P a S e r M e t P r o A r g L e u P r o T h r	460
Db	2009	GGAGGACCTCAGGGTTCCTGAGATTCTCTGACGTGAATTCATGCCCGCTCCAAAG	2068
Qy	461	A S P L e u A S P V a G L u G I P r o T r P H e A r G H S T r A S P P H e A r G I n S e r C y S T r P V a	480
Db	2069	GACCTGACGTAGAGGGCCCTTGCTCCGCCATTATGATTTCAGACAGAGCTGCTGGTTC	2128
Qy	481	A r G a l a l e S e r G I n G L u A S P G I n L e u A P r o C y S T r P G I n a G I u H I S P r o A l a G I u	500
Db	2129	CGTGCCATATCCAGAGAGACCAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTGCCAG	2188
Qy	501	A r G V a l a r S e r A P H e a l a A P r o S e r T h r A S P S e r A S P G I n G L Y T h r P r o P H e a r g	520
Db	2189	CGGTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCGACCAAGGGCACCCCTTCAGA	2248
Qy	521	A A r g A S P G L u G I n P r o C y S A A P r o T h r S e r G I P r o C y S L e u C y S G L Y L e u H I S L e u	540
Db	2249	GCTAGGACGACACAGCCGCTGTGTCCACCTCAGGCGCTTGCTTGCGGCTCCACTTG	2308
Qy	541	G U S e r S e r G I n P H e P r o G I P H e	549
Db	2309	GAAAGTTCTCAGTTCCTCCAGGCTTC	2335
RESULT 12			
AR578515			
LOCUS	AR578515	2072 bp	DNA linear PAT 14-DEC-2004
DEFINITION	Sequence 691 from patent US 6783969.		
ACCESSION	AR578515		
VERSION	AR578515.1	GI:56581311	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2072)		
AUTHORS	Tang,Y.T., Goodrich,R.W., Asundi,V. and Drmanac,R.T.		
TITLE	Cathepsin V-like polypeptides		
JOURNAL	Patent: US 6783969-A 691 31-AUG-2004;		
FEATURES	Nuvelo, Inc.; Sunnyvale, CA		
source	Location/Qualifiers		
	1..2072		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:			
Pred. NO.:	2.36e-89	Length:	2072
Score:	2827.00	Matches:	525
Percent Similarity:	95.6%	Conservative:	0
Best Local Similarity:	95.6%	Mismatches:	2
Query Match:	94.8%	Indels:	22
	6	Gaps:	1
US-10-071-838-2 (1-549) x AR578515 (1-2072)			
Qy	1	M e t A S P V a G L u V a l a G I Y S e r T r P T r P A G I n G L u A r G L u A S P I l e I e M e t	20
Db	163	A T G A C G T G T A G A G T C G G G C A G T T G T G G C A C A G A G C A G A G A C A T C A T T A T G	222
Qy	21	L Y S T r G L u L Y S G I Y H I A r G A l a G L Y L e u P r o G I u A S P L Y S G I P r o L Y S P r o P H e A r g	40
Db	223	A A A T A C G A A A A G G A C C C G A G C T G G G C T G C C A G A G A C A A G G G C C T A G C C T T T C G A	282
Qy	41	S e r T r A S A S A S A S A S A S P H I S L e u G I Y I l e V a H I S G L u T h r G L u L e u P r o P H e u	60
Db	283	A G C T A C A C A C A C G T C G A T C A T T T G G G A T T G T A C A T G A G A C G A G C T G C C T C T G	342
Qy	61	T h r A A r G L u A l a Y S G I l e A r G A r G L u I S e r A r G L Y S e r L Y S T r P V a A S P	80
Db	343	A C T G C G G G A G G C G A G C A A T T C G G C G G A G A T C A G C C G A A A A G A G C A A G T G G T G A T	402
Qy	81	M e L e u G I A S P T r P G I u L Y S T r L Y S S e r S e r A r G L Y S L e u I e A S P A r G A l a T r L Y S	100

Db	403	ATGCTGGAGACTGGGAGAAATACAAAAACGACGAAAGCTCATAGATCGAGCGGTACAAG	462
QY	101	GIYMetProMetAsnIleArgGlyProMetTrpSerValIleuLeuAsnIleGIUmet	120
Db	463	GGAATGCCCATGAAACATCCGGGGCCGAGTGTGTCAGTCCCTCTGAACA CTGAGGAATG	522
QY	121	LYSLeuLYSAsnProGLYArgTYrGlnIleMetLYSGlULYSGLYLYSArgSerSerGIU	140
Db	523	AAGTTGAAAAAACCCCGAAGATATCCAGATCATGAAGAGAGAGGCAAGGTCATCTGAG	582
QY	141	HISIleGlnArgIleAspArgAspValSerGlyThrLeuArgLYHisIlePhePheArg	160
Db	583	CACATCCAGCGCATTCGACCGGGACGTAAAGCGGACATTAAAGAACATATATCTTCAGG	642
QY	161	ASPArgTYrGlyThrLYSGlnArgGluLeuLeuHISIleLeuLeuAlaTYrGIUglUTyr	180
Db	643	GATCGATACGGAACCAAGCAGCGGGAAGTACTCCACATCCTCTGTCATATGAGAGATAT	702
QY	181	ASNProGluValGIlyTYrCysArgAspLeuSerHISIleAlaIleuPheLeuLeuTYr	200
Db	703	AACCCGGAGGTGGCTACTGACGGGACTGTAGCCACATCGCCGCTTGTCTCTCTAT	762
QY	201	LEUProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHISer	220
Db	763	CTTCCTGAGGAGATGATTCTGGGCACTGTGACGTGTCAGCTGTGGCAGTAGAGGCACCTC	822
QY	221	LEUGlnGlyPheHISerProAsnGlyGIlyThrValGlnGlyLeuGlnAspGlnGlnGlu	240
Db	823	CTGCAGGATTTCA CAGCCCAATGCGGGACCGTCCAGGGGCTCCAAGACCAACAGAG	882
QY	241	HISValValAlaThrSerGlnProLYSthrMetGlyHISGlnAspLYSAspLeuCys	260
Db	883	CATGTGTAGCCACGTCAACAACCAAGCATGGGGCATCAG-----	924
QY	261	GLYGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly	280
Db	925	-----ATCTCTCTGGG	936
QY	281	LEUThrLeuArgLeuTrpAspValTYrLeuValGIUglYglGlnAlaLeuMetProIle	300
Db	937	CTCACCCCTGCGCTGTGGGACGTGTACTGTGTAAGAAGCGAACGCGCTTGATGCCGATA	996
QY	301	THRArgIleAlaPheLYSValGlnGlnLYSArgLeuThrLYSThrSerArgCysGlyPro	320
Db	997	ACAAGAATCGCCTTTAAGTTCAAGCAGAACCGCCTCACGAAGCCTCCAGGTGTGCCCG	1056
QY	321	TRPAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu	340
Db	1057	TGGGCACGTTTTTGCAACCGGTTCTGTATACCTGGCCAGGAGTAGAGACACTGTGCTC	1116
QY	341	LYSHISLeuArgAlaSerMetLYSLeuLeuThrArgLYSGlnGlyAspLeuProProPro	360
Db	1117	AAGCATCTTAGGGCCTCTATGAAGAACTAACAGAAGACAGGGGGA CTTGCAACCCCA	1176
QY	361	ALAlysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLYs	380
Db	1177	GCCAAACCCGACGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCACGTGGCGGGAAG	1236
QY	381	THRLeuCysLYSGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1237	ACCCTCTGCAAGGGGACAGCAGCGGCCCTCCAGGCCACCAAGCCCGGTTC CCGCGGCC	1296
QY	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGIlyAla	420
Db	1297	ATTGTGTACGCTTCCCGCCACGCGCACTCGTTCTTCCACACCTGTCTGTGTGGGCT	1356
QY	421	ValArgGluAspThrTYrProValGIlyThrGlnGlyValProSerProAlaLeuAlaGln	440
Db	1357	GTCGGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCACCCCGGCTGTCTCAG	1416
QY	441	GIlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	1417	GGAGGACCTCAGGGGTTCTCTGAGATTCTGCAGTGGAATTCATGCCCGCTCCCAAG	1476

QY 461 AsbLeuAspValGluGlyProTrrPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
|||||
Db 1477 GACCTGGACGTAGAGGCCCTTGTTCCGCCATATGATTTCAGACAGAGCTGTGGGTC 1536

QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
|||||
Db 1537 CGTGCATATCCACAGAGGACCAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTGCGAG 1596

QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
|||||
Db 1597 CCGGTGAGATCGGCTTTCGCTGCACCCACAGCACTGATTCGACACAGGCAACCCCTTCAGA 1656

QY 521 AlaArgAspGluGlnProCysAlaProThxSerGlyProCysLeuCysGlyLeuHisLeu 540
|||||
Db 1657 GCTAGGAGACGAACAGCGGTGTGCTCCCACTCAGGCGCTTGCTGCGGCTCCACTTG 1716

QY 541 GluSerSerGlnPheProProGlyPhe 549
|||||
Db 1717 GAAAGTCTCAGTTCCTCCAGGCTTC 1743

RESULT 13
AX775949 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775949 Sequence 219 from Patent WO03048202.
DEFINITION AX775949
ACCESSION AX775949
VERSION AX775949.1 GI:32693667
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-KB activating gene
JOURNAL Patent: WO 03048202-A 219 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
Source Location/Qualifiers
1.2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB11674.1"
/db_xref="GI:32693668"
CDS
/translation="MDVVEVAGSMWAQERDIIIMKYEKGRAGLPEDKPKPRSSYN
NVDHLGIHETELPPLTAREAKQIRREISRSKMWMDLGDWEKYKSRKLDRAVKGM
PMNIRGPMWSVLNTEEMKLNPRGYQIMKEKGKRSSEHIQRIDRVSGTLAKHIFR
DRYGTQORELHILLAYEENPEVGYCRDLSHIALFLYLPEEDAFWALVOLLASER
HSLQGFHSBNGGTVOGLQDQEHVATSQPKTMGHQISLGLTLRLMDVLYVEGQALM
PIIRIAFYQQKRLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLRASMKULTRKGD
LQPPAKPEQSSASRPVPASRGKITLCKGDROAPPGPAPRPPIWSASPPRAPRSST
PCPGAVREDTYPVGTQGVSPALAQGGPQGSWRFLOMNSMRLPTDLDEGPMFRHY
DFRQSCWVRAISOEDQLAPCWQAEHPAERVRSAPAAPSTDSDQGTFRARDEQPCAPT
SGPCLCGLHLESSQFPFPGF"

ORIGIN
Alignment Scores:
Pred. No.: 2.99e-89 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 95.6% Mismatches: 2
Query Match: 94.8% Indels: 22
DB: 6 Gaps: 1

US-10-071-838-2 (1-549) x AX775949 (1-2647)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
|||||
Db 755 ATGACAGTGTAGAGTCCGCGGCAGTTGTGTGGCAACAAGAGCGAGAGCATCATTTATG 814

QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
|||||
Db 815 AAATACGAAGAAGGACACCGAGCTGGGCTGCCAGAGGACAAGGGCCTTAAGCCTTTCCGA 874

QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
|||||
Db 875 AGCTACAAACAACAGCTGCATCTTGGGGATTGTACATGACAGAGCTGCCCTCTCTG 934

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
|||||
Db 935 ACTGCGCGGAGCGGAGCAAAATTCCGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 994

QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
|||||
Db 995 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAAGAAAGCTCATGATCGAGCGTACAG 1054

QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
|||||
Db 1055 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACACTGAGGAATG 1114

QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
|||||
Db 1115 AAGTTGAAAAACCCCGGAGATACAGATCATGAAGAGAGGCAAGAGTCACTGAG 1174

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
|||||
Db 1175 CACATCCAGCGCATCGACCGGGACGTAAAGCGGACATTAAAGAAACATATATTCTTCAGG 1234

QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
|||||
Db 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTTGCCATATAGAGGTAT 1294

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
|||||
Db 1295 AACCCGAGGTGGGCTACTGCAGGAACTGAGCCACATGCCCGCTTGTCTCTCTAT 1354

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
|||||
Db 1355 CTTCTGAGGAGGATGCATCTGGGCACTGTGCAGCTGTGGCCAGTAGAGGCACTCC 1414

QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
|||||
Db 1415 CTGCAGGGAATTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 1474

QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
|||||
Db 1475 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAG----- 1516

QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
|||||
Db 1517 -----ATCTCTCTCGGG 1528

QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGlnAlaLeuMetProIle 300
|||||
Db 1529 CTCACCGCTGGCCTGTGGACGTGATCTGTAGAAAGGCAACAGCGTTGATGCGGATA 1588

QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
|||||
Db 1589 ACAAGATCGCCTTTAAGGTTCCAGCAGAAGCGCCTCACGAAGACGTCCAGGTGTGGCCG 1648

QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
|||||
Db 1649 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCGCAGGATGAGGACACTGTGCTC 1708

QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
|||||
Db 1709 AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCAACCCCA 1768

QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
|||||
Db 1769 GCCAAACCGGACAAGGTCGTGGCATCCAGGCGCTGTGCGGCTTCAAGTGGCGGGAAG 1828

QY 381 ThrLeuCySlySGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
|||||
Db 1829 ACCCTCTGCAAGGGGAGACAGCAGGCCCTCTCCAGGCCACAGCCCGGTTCCGCGGCC 1888
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCybProGlyGlyAla 420
|||||
Db 1889 ATTTGGTCAGCTTCCCGGCCACCGGCACCTCGTTCTTCCACACACCTGTCTGTGGGCT 1948
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
|||||
Db 1949 GTCCGGGAAGACACCTACCTGTGGCACTCAGGGTGTGCCAGGCCCGCTGGCTCAG 2008
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
|||||
Db 2009 GGAGGACCTCAGGGTTCCTGGAGATTCTCTGACGTGAACTCCATGCCCGCTCCCAACG 2068
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
|||||
Db 2069 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
|||||
Db 2129 CGTGCCATATCCCAAGAGACACAGCTGGCCCCCTGTGCGACGGCTGAACACCTGCGGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
|||||
Db 2189 CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCCACCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCySlyLeuHisLeu 540
|||||
Db 2249 GCTAGGAGCAGAACAGCGGTGTGCTCCCACTCAGGGCCTTGCTTGGCGGCTCCACTTG 2308
QY 541 GluSerSerGlnPheProGlyPhe 549
|||||
Db 2309 GAAAGTTCTCAGTTCCTCCCTCAGGCTTC 2335

RESULT 14
AX833552

LOCUS AX833552 2647 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 676 from Patent EPI347046.
ACCESSION AX833552
VERSION AX833552.1 GI:39919687
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
AUTHORS 1
Isogai,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuho,Y.

TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 676 24-SEP-2003;
Research Association for Biotechnology (JP)

FEATURES
source 1. 2647
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2.99e-89 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 95.6% Mismatches: 2
Query Match: 94.8% Indels: 22
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AX833552 (1-2647)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
|||||
Db 755 ATGACGTTGTAGAGGTCCGCGGCAGCTGTGGGCACAAGAGCGAGACATCATTTATG 814
QY 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
|||||
Db 815 AAATACGAAAGAGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTCGA 874
QY 41 SerTyrAsnAsnAsnValAspHisIleLeuGlyIleValHisGlnThrGluLeuProProLeu 60
|||||
Db 875 AGCTACAAACAACAAGTCGATCATTTGGGGAATTGTACATGAGACGGAAGCTGCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80
|||||
Db 935 ACTCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 994
QY 81 MetLeuGlyAspTrpGlnLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
|||||
Db 995 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAGAAAGCTCATAGATCGAGCTACAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlnMet 120
|||||
Db 1055 GGAATGCCCATGAACATCCGGGGCCCATGTGTCTCATGCTCTCTGAACAAGGAAATG 1114
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGln 140
|||||
Db 1115 AAGTTGAAAAACCCCGAAGATACCATCATGAAGGAGAAAGGCAAGGTCACTTGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
|||||
Db 1175 CACATCCAGCGCATCGACCGGACGTAAACGGGACATTAAGAACATATATCTTCAGG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGlnLeuLeuHisIleLeuLeuAlaTyrGlnGlnTyr 180
|||||
Db 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGATAT 1294
QY 181 AsnProGlnValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
|||||
Db 1295 AACCCGGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 1354
QY 201 LeuProGlnGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
|||||
Db 1355 CTTCCTGAGGAGGATGATTCTGGGCACTGTGACGTGCTGGCCAGTAGAGGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGln 240
|||||
Db 1415 CTGACGGGATTTCACAGCCCAATGCGCGGACCGTCCAGGGGCTCCAAAGACCAACAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
|||||
Db 1475 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAG-----ATCTCTCTCGGG 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
|||||
Db 1517 -----ATCTCTCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
|||||
Db 1529 CTCACCTCGCGCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGCGGTGATGCCGATA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
|||||
Db 1589 ACAAGAAATCGCCTTTAAGGTTCAAGAAAGCGCCTCAGAAAGCGTCCAGGTGTGCCCG 1648
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
|||||
Db 1649 TGAGCAAGTTTTCGAACCGGTTCTGTATACCTGGGCAAGGATGAGACACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
|||||
Db 1709 AAGCATTTAAGGGCCTCTATGAAGAACTAACAAAGAAAGCAAGGGGACCTGCCAACCCCA 1768
QY 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380

|||||
Db 1769 GCCAACCAGCAAGGGTCGTCCGCATCCAGGCGCTGTGCCGCTTCAAGTGGCGGAG 1828
Qy 381 ThrLeuCyblyeglyAsparglnAlaProProglyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGCAAGGGGAGCAGGAGGCCCTTCCAGGCCACCAGCCCGGTTCCCGCGGCC 1888
Qy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCySPROglyAla 420
Db 1889 ATTGGTCAGCTTCCCGCCACGCGGCACCTCTCTTCCACACACCTGTCTGTGGGCT 1948
Qy 421 ValArgGluAspThrTyProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCCGGAGAGACACCTAACCTGTGGGCATCAGGGTGTGCCACCGCCGCTGTCTAG 2008
Qy 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 2009 GGAGGACCTCAGGGTCTCGAGATTCCTGCAGTGAACCTCCATGCCCGCCTCCCAACG 2068
Qy 461 AspleuAspValGluGlyProTrpPheArgHisTyAspPheArgGlnSerCySTrpVal 480
Db 2069 GACCTGACCGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 2128
Qy 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCySTrpGlnAlaGlnHisProAlaGln 500
Db 2129 CGTGCATATCCAGAGGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTGCGAG 2188
Qy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 2189 CGGGTAGATCGGCTTTCGTGCACCCACCACTGATTCGACACGAGGACACCCCTTCAGA 2248
Qy 521 AlaArgAspGluGlnProCyAlaProThrSerGlyProCybLeuCybGlyLeuHisLeu 540
Db 2249 GCTAGGAGACGACGCGGTGTGCTCCCACTCAGGGCCTTGCTGCGGCGCTCCACTTG 2308
Qy 541 GluSerSerGlnPheProProGlyPhe 549
Db 2309 GAAAGTCTCAGTTCCTCCAGGCTTC 2335
RESULT 15
AK095385 2647 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ38066 fis, clone CTONG2015330, moderately
DEFINITION similar to H.sapiens mRNA for tre oncogene.
ACCESSION AK095385 GI:21754631
VERSION AK095385.1
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
Sekine,M., Oabayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
Yosida,M., Hotta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuki,H., Oshima,A.,
Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,

Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Sato,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE
AUTHORS 2
Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M.,
Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL 3 (bases 1 to 2647)
REFERENCE Isogai,T. and Yamamoto,J.
AUTHORS Direct Submission
TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source Location/Qualifiers
1. 2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2015330"
/tissue_type="tongue, tumor tissue"
/clone_fib="CTONG2"
/note="cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 2.99e-89 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.6% Conservative: 0
Best Local Similarity: 95.6% Mismatches: 2
Query Match: 94.8% Indels: 22
DB: 8 Gaps: 1
US-10-071-838-2 (1-549) x AK095385 (1-2647)
Qy 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
Db 755 ATGGACGTGTAGAGGTGCGCGGCAGTTGCTGGGCACAAAGACGAGAGACATCATATG 814
Qy 21 LysTyGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 815 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTAAGCCTTTTGA 874
Qy 41 SerTyAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
Db 875 AGCTACAACAACAACGTGCATCTTGGGAGATTGTACATGAGACGAGCTGCCTCTCG 934
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

Db 935 ACTGCCGGAGCGAAGCAAAATTCCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 994
QY 81 MetLeuGlyAspTrpGluIleYsTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 995 ATGCTGGAGACTGGAGAAATACAAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
Db 1055 GGAATGCCCATGAAACATCCGGGGCCCGATGTGTCAAGTCCCTCTGAACACTGAGGAAATG 1114
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 1115 AAGTTGAAAAACCCCGAAGATACCAGATCATGAAGAGAGGCAAGAGGTCACTTGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAAAGCATATATTCTTCAG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 1235 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCGGCATATGAGAGTAT 1294
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 1295 AACC CGAGGTGGGTACTGACAGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 1355 CTTCCTGAGAGAGATGATTCGTGGCAGTGTGCTGCTGCGCAGTGAAGGCACTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1415 CTGCAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 1475 CATGTGTAGCCACGTCACAACCCAAAGACCATGGGCATCAG----- 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 1517 -----ATCTCTCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
Db 1529 CTCACCTGCGCCTGTGGAGCGTGTATCTGTGTAGAAAGCGGAACAGCGCTTGATGCCGAT 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1589 ACAAGAATCGCTTTAAGTTCAAGCAGAACCGCTCAGCAAGACCTCCAGGTGTGCCCG 1648
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1649 TGGCAGCGTTTGGCAACCGGTTCTGTGATACCTGGGCCAAGGATGAGGACACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1709 AAGCATCTTAGGGCTCTATGAAAGAACTAAAGAAAGCAGGGGGAACCTGCAACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1769 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCCGCTTCACTGTGCGGGAAG 1828
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGCAAGGGGAGCAGGCAAGGCCCTCCAGGCCACCGGTTCCCGCGGCC 1888
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1889 ATTTGGTCAGCTTCCCGCCACGCGGCACTCTTCCACACCTGTCTGTGTGGGCT 1948
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCCGGGAAGACCTACCTGTGGGACTCAGGGGTGTGCCACGCCCGCCCTGTGCTCAG 2008

QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 2009 GGAGGACCTCAGGGTTCTGTGAGATTCTGCAGTGAAGTCCATGCCCGCTCCCAAGC 2068
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db 2069 GACCTGACGTAAGGGGCCCTTGGTTCGCGCATATGATTTGACAGAGCTGTGGGTC 2128
QY 481 ArgAlaIleSerGlnLysAspGlnLeuAlaProCysTrpGlnAlaGluHisProAlaGlu 500
Db 2129 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGCGAGGCTGAACACCTCGGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 2189 CGGTGAGATGGCTTTCGCTGCAACCCACATGATTCGACCAAGGCAACCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 2249 GCTAGGACGAAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTGTGCGGCTTCACCTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 2309 GAAATCTCAGTTCCCTCCAGGCTTC 2335

Search completed: April 5, 2006, 16:34:43
Job time : 6744 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 14:01:23 ; Search time 1619 Seconds
(without alignments)
10031.537 Million cell updates/sec

Title: US-10-071-838-1

Perfect score: 1964

Sequence: 1 atggacgtgtagaggtcgc.....aaatgttctctgtgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	1964	100.0	1964	US-10-071-838-1	Sequence 1, Appl1
2	1929.8	98.3	1993	US-10-839-882-31	Sequence 31, Appl1
3	1836.4	93.5	2084	US-10-302-172-692	Sequence 692, Appl1
4	1812.4	92.3	2072	US-10-302-172-691	Sequence 691, Appl1
5	1807.4	92.0	2647	US-10-108-260A-676	Sequence 676, Appl1
6	1794	91.3	2146	US-10-071-838-3	Sequence 3, Appl1
7	1750	89.1	1862	US-10-071-838-5	Sequence 5, Appl1
8	1638.8	83.4	1752	US-10-094-466-61	Sequence 61, Appl1
9	1220	62.1	7856	US-10-450-763-21335	Sequence 21335, A
10	1220	62.1	7856	US-10-450-763-24017	Sequence 24017, A
11	1220	62.1	7878	US-09-962-436-592	Sequence 562, Appl1
12	1220	62.1	7878	US-10-843-641A-3021	Sequence 3021, Appl1
13	1220	62.1	7878	US-10-956-157-1259	Sequence 1259, Appl1
14	999	50.9	8180	US-10-450-763-24016	Sequence 24016, A
15	999	50.9	8201	US-10-285-975A-272	Sequence 272, Appl1
16	999	50.9	8201	US-10-278-698-640	Sequence 640, Appl1
17	999	50.9	8201	US-10-278-698-1037	Sequence 1037, Appl1
18	999	50.9	8284	US-10-450-763-21339	Sequence 21339, A
19	999	50.9	8408	US-10-098-841-51	Sequence 51, Appl1
20	999	50.9	8420	US-10-098-841-50	Sequence 50, Appl1
21	882.8	44.9	9804	US-09-764-869-1829	Sequence 1829, Appl1
22	882.8	44.9	9804	US-09-764-891-7861	Sequence 7861, Appl1
23	882.8	44.9	9804	US-10-091-504-1829	Sequence 1829, Appl1

24	882.8	44.9	9804	US-10-227-577-1829	Sequence 1829, Appl1
25	879.6	44.8	9404	US-09-764-869-1826	Sequence 1826, Appl1
26	879.6	44.8	9404	US-09-764-891-7859	Sequence 7859, Appl1
27	879.6	44.8	9404	US-10-091-504-1826	Sequence 1826, Appl1
28	879.6	44.8	9404	US-10-227-577-1826	Sequence 1826, Appl1
29	878	44.7	886	US-09-764-869-1832	Sequence 1832, Appl1
30	878	44.7	886	US-10-091-504-1832	Sequence 1832, Appl1
31	878	44.7	886	US-10-227-577-1832	Sequence 1832, Appl1
32	876.4	44.6	886	US-09-764-869-1831	Sequence 1831, Appl1
33	876.4	44.6	886	US-10-091-504-1831	Sequence 1831, Appl1
34	876.4	44.6	886	US-10-227-577-1831	Sequence 1831, Appl1
35	876.4	44.6	6696	US-09-764-869-1821	Sequence 1821, Appl1
36	876.4	44.6	6696	US-09-764-891-7856	Sequence 7856, Appl1
37	876.4	44.6	6696	US-10-091-504-1821	Sequence 1821, Appl1
38	876.4	44.6	6696	US-10-227-577-1821	Sequence 1821, Appl1
39	876.4	44.6	9805	US-09-764-869-1824	Sequence 1824, Appl1
40	876.4	44.6	9805	US-09-764-891-7857	Sequence 7857, Appl1
41	876.4	44.6	9805	US-10-091-504-1824	Sequence 1824, Appl1
42	876.4	44.6	9805	US-10-227-577-1824	Sequence 1824, Appl1
43	864.4	44.0	9796	US-09-764-869-1825	Sequence 1825, Appl1
44	864.4	44.0	9796	US-09-764-891-7858	Sequence 7858, Appl1
45	864.4	44.0	9796	US-10-091-504-1825	Sequence 1825, Appl1

ALIGNMENTS

RESULT 1					
US-10-071-838-1					
; Sequence 1, Application US/10071838					
; Publication No. US20030044814A1					
; GENERAL INFORMATION:					
; APPLICANT: Li, Jing					
; APPLICANT: Powers, Scott					
; APPLICANT: Xiang, Phil					
; APPLICANT: Peng, Yue					
; APPLICANT: Tularek Inc.					
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene					
; FILE REFERENCE: 018781-002610US					
; CURRENT APPLICATION NUMBER: US/10/071,838					
; CURRENT FILING DATE: 2002-05-07					
; PRIOR APPLICATION NUMBER: US 60/267,615					
; PRIOR FILING DATE: 2001-02-08					
; NUMBER OF SEQ ID NOS: 18					
; SOFTWARE: Patentin Ver. 2.1					
; SEQ ID NO 1					
; LENGTH: 1964					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
; FEATURE:					
; OTHER INFORMATION: human PRC17					
; FEATURE:					
; NAME/KEY: CDS					
; LOCATION: (1)..(1650)					
; US-10-071-838-1					
Query Match 100.0%; Score 1964; DB 5; Length 1964;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	ATGACGCTGTAGAGTTCGGGCGAGTTGGTGCCACAGAGCGGAGGACATCATATG	60		
DB	1	ATGACGCTGTAGAGTTCGGGCGAGTTGGTGCCACAGAGCGGAGGACATCATATG	60		
QY	61	AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA	120		
DB	61	AAATACGAAAGGACACCGAGCTGGCTGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA	120		
QY	121	AGCTACAAACAACAAGCTGCATCTTTGGGATTTGTACATGAGACGAGCTGCTCTCTG	180		
DB	121	AGCTACAAACAACAAGCTGCATCTTTGGGATTTGTACATGAGACGAGCTGCTCTCTG	180		
QY	181	ACTGCGCGGAGGCGAAGCAATTTCGGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT	240		

```
Db 181 ACTGCGGGAGGCGAAGCAAAATTCGCGGGAGATCAGCCGAAAGAGCAAGTGGGTGAT 240
Qy 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAAAGTCTATAGATCGAGCGTACAAG 300
Db 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAGAAAGTCTATAGATCGAGCGTACAAG 300
Qy 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAAACATTGAGGAAATG 360
Db 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAAACATTGAGGAAATG 360
Qy 361 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGGAGAAAGGCAAGAGTCACTGAG 420
Db 361 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGGAGAAAGGCAAGAGTCACTGAG 420
Qy 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGAAAGCATATATTCTTCAGG 480
Db 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGAAAGCATATATTCTTCAGG 480
Qy 481 GATGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAT 540
Db 481 GATGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAT 540
Qy 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAAGCACAATCGCCGCTGTCTCTCTAT 600
Db 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAAGCACAATCGCCGCTGTCTCTCTAT 600
Qy 601 CTTCTGAGAGAGATGATCTGTGGGACCTGTGCAAGTGTGGCCAGTGAAGAGCACTCC 660
Db 601 CTTCTGAGAGAGATGATCTGTGGGACCTGTGCAAGTGTGGCCAGTGAAGAGCACTCC 660
Qy 661 CTGCAAGGATTTCACAGCCCAAAATGCGGGAACCGTCCAGGGGCTCAAGACCAAGAGAG 720
Db 661 CTGCAAGGATTTCACAGCCCAAAATGCGGGAACCGTCCAGGGGCTCAAGACCAAGAGAG 720
Qy 721 CATGTGTAGCCACGTCAACAACCAAGACCATGAGGACATGAGCAAGAAAGATCTATGT 780
Db 721 CATGTGTAGCCACGTCAACAACCAAGACCATGAGGACATGAGCAAGAAAGATCTATGT 780
Qy 781 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGAATTGACGGATCTCTCGGG 840
Db 781 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGAATTGACGGATCTCTCGGG 840
Qy 841 CTCACCTGCGGCTGTGGGACGTGTATCTGTGTAAGAGGCAAGGCGTTGATGCCGATA 900
Db 841 CTCACCTGCGGCTGTGTGGACGTGTATCTGTGTAAGAGGCAAGGCGTTGATGCCGATA 900
Qy 901 ACAAGATCGCCTTTAAGGTTACAGCAAGCGCCTCAAGCAAGCATCCAGGTGCGCCG 960
Db 901 ACAAGATCGCCTTTAAGGTTACAGCAAGCGCCTCAAGCAAGCATCCAGGTGCGCCG 960
Qy 961 TGGGACGTTTTTGCAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC 1020
Db 961 TGGGACGTTTTTGCAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC 1020
Qy 1021 AAGCATCTTAGGGCCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1021 AAGCATCTTAGGGCCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1080
Qy 1081 GCCAAACCCGAGCAAGGGTCTCGGCATCCAGGCTGTGCGGCTTCAAGTGGGCGGAAG 1140
Db 1081 GCCAAACCCGAGCAAGGGTCTCGGCATCCAGGCTGTGCGGCTTCAAGTGGGCGGAAG 1140
Qy 1141 ACCCTCTGCAAGGGGACAGGAGCGCCTCCAGGCGCCACAGCCCGGTTCCCGGCGCC 1200
Db 1141 ACCCTCTGCAAGGGGACAGGAGCGCCTCCAGGCGCCACAGCCCGGTTCCCGGCGCC 1200
Qy 1201 ATTTGTAGCTTCCCGCACGGGACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1260
Db 1201 ATTTGTAGCTTCCCGCACGGGACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1260
Qy 1261 GTCCGGGAAGACACTACCTGTGGGCACTCAGGTTGTGCCAGCCCGCCTGCTCAG 1320
Db 1261 GTCCGGGAAGACACTACCTGTGGGCACTCAGGTTGTGCCAGCCCGCCTGCTCAG 1320
```

```
Db 1261 GTCCGGGAAGACACTACCTGTGGGCACTCAGGTTGTGCCAGCCCGCCTGCTCAG 1320
Qy 1321 GGAGACCTCAGGGTTCCTGAGAGATTCTTGCAGTGGAATCCATGCCCCGCTCCCAACG 1380
Db 1321 GGAGACCTCAGGGTTCCTGAGAGATTCTTGCAGTGGAATCCATGCCCCGCTCCCAACG 1380
Qy 1381 GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTAATGATTTCAGACAGAGCTGTGGGTC 1440
Db 1381 GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTAATGATTTCAGACAGAGCTGTGGGTC 1440
Qy 1441 CGTGCCATATCCAGAGAGACCAAGCTGAGCCCTGCTGCGAGGCTGAACAACCTGCGAG 1500
Db 1441 CGTGCCATATCCAGAGAGACCAAGCTGAGCCCTGCTGCGAGGCTGAACAACCTGCGAG 1500
Qy 1501 CGGGTGAAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCACCCCCTCAGA 1560
Db 1501 CGGGTGAAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCACCCCCTCAGA 1560
Qy 1561 GCTAGGGAAGAACAGACCGGTGTGCTCCCACTTCAGAGGCTTGCCCTGCGGCTCCACTTG 1620
Db 1561 GCTAGGGAAGAACAGACCGGTGTGCTCCCACTTCAGAGGCTTGCCCTGCGGCTCCACTTG 1620
Qy 1621 GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGCTCATGCGTGAT 1680
Db 1621 GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGCTCATGCGTGAT 1680
Qy 1681 AATTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCTCGTTTATTTTGGTTAA 1740
Db 1681 AATTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCTCGTTTATTTTGGTTAA 1740
Qy 1741 CTTATGAATAATGTATTAGAAGAGTGCAGCTGCAGAGAGATTCAAGATGGAACACACC 1800
Db 1741 CTTATGAATAATGTATTAGAAGAGTGCAGCTGCAGAGAGATTCAAGATGGAACACACC 1800
Qy 1801 AGACCCCAAGATCACAAAAGCCAACCATGCCAGCCCTCCAGACAACCCCAAGCAGAC 1860
Db 1801 AGACCCCAAGATCACAAAAGCCAACCATGCCAGCCCTCCAGACAACCCCAAGCAGAC 1860
Qy 1861 CATGCTTCTGAATTCTGACGACACCGGTGAGCTGCTTGTACTTTAACTCATGGAAG 1920
Db 1861 CATGCTTCTGAATTCTGACGACACCGGTGAGCTGCTTGTACTTTAACTCATGGAAG 1920
Qy 1921 ATAACTACCTTCACGTTTGAATAAATGTTCTGTGAATG 1964
Db 1921 ATAACTACCTTCACGTTTGAATAAATGTTCTGTGAATG 1964

RESULT 2
US-10-839-882-31
; Sequence 31, Application US/10839882
; Publication No. US20040203106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: CORLEY, Neil C.
; APPLICANT: LAL, Preeti
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: JUNMING, Yang
; APPLICANT: SHIH, Leo L.
; TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
; FILE REFERENCE: PF-0619 PCT
; CURRENT APPLICATION NUMBER: US/10/839,882
; CURRENT FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US/09/807,452
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/175,737; unassigned; 60/118,559; 09/249,740; unassigned
; 60/154,336
; PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11;
; 1999-04-22
```



```

; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 31
; LENGTH: 1993
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 4325626CB1
US-10-839-882-31

```

Query Match	98.3%	Score 1929.8;	DB 8;	Length 1993;
Best Local Similarity	99.6%	Pred. No. 0;		
Matches 1945; Conservative	0;	Mismatches	7;	Indels 1; Gaps 1;

QY	1	ATGACGTGTAGAGGTCCGGGCGAGTTGTGGGCA	CAAGAGCGAGACATCATATG	60
Db	42	ATGACGTGTAGAGTCCGGGCGAGTTGTGGGCA	CAAGAGCGAGACATCATATG	101
QY	61	AAATACGAAAAAGGCA	CCGAGCTGGCTGCCAGAGCAAGGGGCTAAGCCTTTTGA	120
Db	102	AAATACGAAAAAGGCA	CCGAGCTGGCTGCCAGAGCAAGGGGCTAAGCCTTTTGA	161
QY	121	AGCTACAACAACA	CGTCGATCATTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG	180
Db	162	AGCTACAACAACA	CGTCGATCATTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG	221
QY	181	ACTCGCGGGAGGCGAAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGGAT		240
Db	222	ACTCGCGGGAGGCGAAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGGAT		281
QY	241	ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAG		300
Db	282	ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAG		341
QY	301	GGAATGCCCATGAACATCCGGGGCCCGAATGTGCTCAGTCCCTCTGAACATTGAGGAATG		360
Db	342	GGAATGCCCATGAACATCCGGGGCCCGAATGTGCTCAGTCCCTCTGAACATTGAGGAATG		401
QY	361	AAGTGAAAAACCCCGGAAGATACCAGATCATGAAGGAGAAAGGCAAGGTCATCTGAG		420
Db	402	AAGTGAAAAACCCCGGAAGATACCAGATCATGAAGGAGAAAGGTCATCTGAG		461
QY	421	CACATCCAGCGCATCGACCGGAGCTGAAGCGGACATTAAAGAACATATATTCTTCAGG		480
Db	462	CACATCCAGCGCATCGACCGGAGCTGAAGCGGACATTAAAGAACATATATTCTTCAGG		521
QY	481	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGGAGTAT		540
Db	522	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGGAGTAT		581
QY	541	AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT		600
Db	582	AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATGCGCGCTTGTCTCTCTAT		641
QY	601	CTTCTGAGAGGATGCAATTCGTGGCACTGTGCACTGTGCGCCAGTGAGAGGCACTCC		660
Db	642	CTTCTGAGAGGATGCAATTCGTGGCACTGTGCACTGTGCGCCAGTGAGAGGCACTCC		701
QY	661	CTGCAAGGATTTCA	CAGCCCAAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG	720
Db	702	CTGCAAGGATTTCA	CAGCCCAAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG	761
QY	721	CATGTGTAGCCA	CGTCAACACCCCAAGACCATGGGGCATCAGSACAAGAAAGATCTATGT	780
Db	762	CATGTGTAGCCA	CGTCAACACCCCAAGACCATGGGGCATCAGSACAAGAAAGATCTATGT	821
QY	781	GGGCAGTGTTCCCGTTAGGCTGCTCATCCGATATTGATTGA	CGGGATCTCTCGGG	840
Db	822	GGGCAGTGTTCCCGTTAGGCTGCTCATCCGATATTGATTGA	CGGGATCTCTCGGG	881
QY	841	CTCACCTCGGCTGTGGAGCGTATACTGTAGAAAGCGAACAGGCGTTGATGCCGATA		900

Db	882	CTCACCCCTGCGCCTGTGGACGTGTAATCTGTAGAAAGCGAACAAGCCTGTGATGCCGATA	941
QY	901	ACAAGAATCGCCTTTAAGGTTTCAGCAGAAAGCGCCTCAGAGACGTCCAGGTGTGCCCG	960
Db	942	ACAAGAATCGCCTTTAAGGTTTCAGCAGAAAGCGCCTCAGAGACGTCCAGGTGTGCCCG	1001
QY	961	TGGGCACGTTTTTTCCAACCGGTTGCTGATACCTGGGCCAGGGATGAGGACATGTGCTC	1020
Db	1002	TGGGCACGTTTTTTCCAACCGGTTGCTGATACCTGGGCCAGGGATGAGGACATGTGCTC	1061
QY	1021	AAGCATCTTAGGGCCTCTATGAAAGAACTAAACAAGAAAGAGGGGACCTGCGCACCCCA	1080
Db	1062	AAGCATCTTAGGGCCTCTATGAAAGAACTAAACAAGAAAGAGGGGACCTGCGCACCCCA	1121
QY	1081	GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCTGTGCGGCTTACGTGGCGGAAG	1140
Db	1122	GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCTGTGCGGCTTACGTGGCGGAAG	1181
QY	1141	ACCCTCTGCAAGGGGACAGGCAAGGCCCTCCAGGCCCAACAGCCCGGTTCCCGCGCCC	1200
Db	1182	ACCCTCTGCAAGGGGACAGGCAAGGCCCTCCAGGCCCAACAGCCCGGTTCCCGCGCCC	1241
QY	1201	ATTGGTCAAGCTTCCCGCCACGGGACCTGTTCTTTCACACCCCTGTCTGTGGGGCT	1260
Db	1242	ATTGGTCAAGCTTCCCGCCACGGGACCTGTTCTTTCACACCCCTGTCTGTGGGGCT	1301
QY	1261	GTCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAAGCCCGCCCTGCTCAG	1320
Db	1302	GTCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAAGCCCGCCCTGCTCAG	1361
QY	1321	GGAGGACCTCAGGGGTTCTGTGAGATTCTCTGAGTGGAACTCCATGCCCGCCTCCAAAG	1380
Db	1362	GGAGGACCTCAGGGGTTCTGTGAGATTCTCTGAGTGGAACTCCATGCCCGCCTCCAAAG	1421
QY	1381	GACCTGACGTAGAGGGCCCTGTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC	1440
Db	1422	GACCTGACGTAGAGGGCCCTGTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC	1481
QY	1441	CGTGCCATATCCACAGAGAGACAGCTGGCCCCCTGTGCGCAGAGCTGAACACCTGCGGAG	1500
Db	1482	CGTGCCATATCCACAGAGAGACAGCTGGCCCCCTGTGCGCAGAGCTGAACACCTGCGGAG	1541
QY	1501	CGGGTGAGATCGGCTTTCGCTGACCCAGCACTGATTCCGACCAAGGGCAACCCCTTCAGA	1560
Db	1542	CGGGTGAGATCGGCTTTCGCTGACCCAGCACTGATTCCGACCAAGGGCAACCCCTTCAGA	1601
QY	1561	GCTAGGGAAGAAAGAGCCGTGTGCTCCCACTCAGGGCCTTGCTGTGCGGCTCCACTTG	1620
Db	1602	GCTAGGGAAGAAAGAGCCGTGTGCTCCCACTCAGGGCCTTGCTGTGCGGCTCCACTTG	1661
QY	1621	GAAAGTTCTCAGTTCCTCCAGAGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT	1680
Db	1662	GAAAGTTCTCAGTTCCTCCAGAGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT	1721
QY	1681	AATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCCGCTCTCGTTTATTTTGGTTAAA	1740
Db	1722	AATTTCCCTAGGCTTAACAACCCCAAGCAAG- TTCCGATCTCTCGTTTATTTTGGTTAAA	1780
QY	1741	CTTATGAAAAATGTATTAAAGAAAGGTGAGCTTCAGAGAGATTTCAGAGATGGAACACACC	1800
Db	1781	CTTATGAAAAATGTATTAAAGAAAGGTGAGCTTCAGAGAGATTTCAGAGATGGAACACACC	1840
QY	1801	AGACCCGAGATCACAAGGCAACCATGCCAGGCCCTCCAGCAACCCCGAGCCCAAGAC	1860
Db	1841	AGACCCGAGATCACAAGGCAACCATGCCAGGCCCTCCAGCAACCCCGAGCCCAAGAC	1900
QY	1861	CATCGTTCTGAATTCTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG	1920
Db	1901	CATCGTTCTGAATTCTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG	1960
QY	1921	ATAACTACCTTCACGTTTTGAAATAAATGTTTC	1953
Db	1961	ATAACTACCTTCATGTTTTGAAATAAATGAGGTC	1993

```
RESULT 3
US-10-302-172-692
; Sequence 692, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an
; FILE REFERENCE: 803_1CNCB
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-08-20
; PRIOR FILING DATE: 2002-08-20
; PRIOR FILING DATE: 2002-03-05
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc_fl_genes Version 2.0
; SEQ ID NO 692
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)..(1755)
US-10-302-172-692
```

```
Query Match          93.5%; Score 1836.4; DB 7; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;
```

```
QY 1 ATGAGCGTGTAGAGGTGCGGGCAGTTGGTGGCCACAAGAGCGAGACATCATATG 60
DB 163 ATGAGCGTGTAGAGGTGCGGGCAGTTGGTGGCCACAAGAGCGAGACATCATATG 222
QY 61 AAATACGAAAGGAGACACCGAGCTGGGCTGCAGAGACAAAGGGCCTAAGCCTTTTGA 120
DB 223 AAATACGAAAGGAGACACCGAGCTGGGCTGCAGAGACAAAGGGCCTAAGCCTTTTGA 282
QY 121 AGCTACAACAACAGCTCGATCATTTGGGGATTGTACATGAGACGGAGCTGCTCTCTG 180
DB 283 AGCTACAACAACAGCTCGATCATTTGGGGATTGTACATGAGACGGAGCTGCTCTCTG 342
QY 181 ACTGCGCGGGAGCGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 240
DB 343 ACTGCGCGGGAGCGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 402
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAGACAGAGAAAGCTCATAGATCGAGCGTACAAG 300
DB 403 ATGCTGGGAGACTGGGAGAAATACAAAGACAGAGAAAGCTCATAGATCGAGCGTACAAG 462
QY 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATG 360
DB 463 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATG 522
QY 361 AAGTTGAAAAAAGCCCGGAAGATACCAAGATCATGAAGAGAGGCAAGAGTCAATGAG 420
DB 523 AAGTTGAA----- 529
QY 421 CACATCCAGCGCATCGACCGGGAGCGTAAAGCGGGACATTAAAGAACATATATTCTTCAAG 480
DB 530 -ACATCCAGCGCATCGACCGGGAGCGTAAAGCGGGACATTAAAGAACATATATTCTTCAAG 588
QY 481 GATCGATACGGAACCAAGCAGCGGGAACTACTCCACATCTCTCTGGCATATGAGGAGTAT 540
DB 589 GATCGATACGGAACCAAGCAGCGGGAACTACTCCACATCTCTCTGGCATATGAGGAGTAT 648
QY 541 AACCCGAGAGTGGGCTACTGCAAGGAGCACTGAGCCACATCGCGCTGTCTCTCTAT 600
```

```
DB 649 AACCCGAGAGTGGGCTACTGCAAGGAGCACTGAGCCACATCGCCGCTTGTCTCTCTAT 708
QY 601 CTTCTGAGAGAGATGATCTTGGGCACTGTGACGCTGTGGCAGTGAAGGCACTCC 660
DB 709 CTTCTGAGAGAGATGATCTTGGGCACTGTGACGCTGTGGCAGTGAAGGCACTCC 768
QY 661 CTGCAAGGATTTACAGAGCCCAAATGGCCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
DB 769 CTGCAAGGATTTACAGAGCCCAAATGGCCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 828
QY 721 CATGTGTAGCCACGTACAAACCCAAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 780
DB 829 CATGTGTAGCCACGTACAAACCCAAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 888
QY 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
DB 889 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 948
QY 841 CTCACCTGCGCCTGTGGAGGTGTATCTGTAGAGGCGAAACAGGCGTTGATGCCGATA 900
DB 949 CTCACCTGCGCCTGTGGAGGTGTATCTGTAGAGGCGAAACAGGCGTTGATGCCGATA 1008
QY 901 ACAAGATCGCCTTTAAGTTACAGAGAGCGCCTCAAGAGACGTCCAGGTGCCCCG 960
DB 1009 ACAAGATCGCCTTTAAGTTACAGAGAGCGCCTCAAGAGACGTCCAGGTGCCCCG 1068
QY 961 TGGGCACTTTTTCGAACCGGTTCTGTATACCTGGGCCAGGGATGAGGACACTGTGCTC 1020
DB 1069 TGGGCACTTTTTCGAACCGGTTCTGTATACCTGGGCCAGGGATGAGGACACTGTGCTC 1128
QY 1021 AAGCATCTTAGGGCCTCTATGAGAACTAACAAAGAAAGCAGGGGACCTGCGACCCCA 1080
DB 1129 AAGCATCTTAGGGCCTCTATGAGAACTAACAAAGAAAGCAGGGGACCTGCAACCCCA 1188
QY 1081 GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCGCTGTGCCGCTTCACTGGCGGGAAG 1140
DB 1189 GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCGCTGTGCCGCTTCACTGGCGGGAAG 1248
QY 1141 ACCCTGCAAGGGGACAGGAGGCCCCCTCCAGGCCCAAGCCCCGTTCCGCGGCC 1200
DB 1249 ACCCTGCAAGGGGACAGGAGGCCCCCTCCAGGCCCAAGCCCCGTTCCGCGGCC 1308
QY 1201 ATTTGTCAGCTTCCCGCCACCGGCACTCGTTCTTCCAACACCTGTCTGTGGGGCT 1260
DB 1309 ATTTGTCAGCTTCCCGCCACCGGCACTCGTTCTTCCAACACCTGTCTGTGGGGCT 1368
QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGCCCAAGCCCCGCTGCTCAG 1320
DB 1369 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGCCCAAGCCCCGCTGCTCAG 1428
QY 1321 GGAGACCTCAGGGTCTGTGAGATTCTGCAATGGAATCCATGCCCGCTCCCAAG 1380
DB 1429 GGAGACCTCAGGGTCTGTGAGATTCTGCAATGGAATCCATGCCCGCTCCCAAG 1488
QY 1381 GACCTGACGTAGAGGCGCTTGGTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 1440
DB 1489 GACCTGACGTAGAGGCGCTTGGTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 1548
QY 1441 CGTGCAATATCCAGAGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1500
DB 1549 CGTGCAATATCCAGAGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1608
QY 1501 CGGGTGAATCGGCTTTCGTGACCAACCAAGCATGATTCGACCAAGGCAACCCCTTCAAG 1560
DB 1609 CGGGTGAATCGGCTTTCGTGACCAACCAAGCATGATTCGACCAAGGCAACCCCTTCAAG 1668
QY 1561 GCTAGGAGCAACAGCGGTGTGCTCCACCTCAGGGGCTTGCCTGTGCGGCTTCACTTG 1620
DB 1669 GCTAGGAGCAACAGCGGTGTGCTCCACCTCAGGGGCTTGCCTGTGCGGCTTCACTTG 1728
QY 1621 GAAAGTCTCAGTTCCTCAGGCTTCTAGAAGCATCTGGGCCAGGCTCATGCTGAT 1680
```

Db 1729 GAAAGTTCAGTTCCTCCAGGCTTCTAGAACATCTGGGCCAGGGCTCATGCGTGAT 1788
QY 1681 AATTTCCCTAGGCTTAAACAACCCCAAGCAAGCTTGCCTCGTTTATTTTGGTTAA 1740
Db 1789 AATTTCCCTAGGCTTAAACAACCCCAAGCAAGCTTGCATCCTCGTTTATTTTGGTTAA 1848
QY 1741 CTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGATTCAGAGATGGAACACACC 1800
Db 1849 CTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGATTCAGAGATGGAACACACC 1908
QY 1801 AGACCCCAAGATCAACAAGCCCAACCATGCCCCCTCCAGACACCCCAAGCCCAAGC 1860
Db 1909 AGACCCCAAGATCAACAAGCCCAACCATGCCCCCTCCAGACACCCCAAGCCCAAGC 1968
QY 1861 CATCGTCTGAAATTCAGACACACCGTGAGCCTGCTTGTACTTTAAACTCATGAAG 1920
Db 1969 CATCGTCTGAAATTCAGACACACCGTGAGCCTGCTTGTACTTTCAAACTCATGAAG 2028
QY 1921 ATAACCTACCTTCACGTTTGAATAATGTTCTCTGTGAAATG 1964
Db 2029 ATAACCACTTCATGTTTGAATAATGTTCTCTGTGAAATG 2072

RESULT 4

US-10-302-172-691
; Sequence 691, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20040053250a1el Arginine-rich Protein-like Nucleic Acids an
; FILE REFERENCE: 803_1CNP
; CURRENT APPLICATION NUMBER: US/10/302,172
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 691
; LENGTH: 2072
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)..(1743)
US-10-302-172-691

Query Match 92.3%; Score 1812.4; DB 7; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY 1 ATGGACGTGTAGAGGTGCGGGGCACTGTGGGCACAAGACGAGAGACATCATTTATG 60
Db 163 ATGGACGTGTAGAGGTGCGGGGCACTGTGGGCACAAGACGAGAGACATCATTTATG 222
QY 61 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAGGGGCTTAAGCCTTTTGA 120
Db 223 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGACAAGGGGCTTAAGCCTTTTGA 282
QY 121 AGCTACAACAACAAGCTGATCATTTGGGGATTGTACATGACGGAAGTGCCTCTCTG 180
Db 283 AGCTACAACAACAAGCTGATCATTTGGGGATTGTACATGACGGAAGTGCCTCTCTG 342
QY 181 ACTGCGGGGAGCGGAAGCAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGGTGAT 240
Db 343 ACTGCGGGGAGCGGAAGCAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGGTGAT 402

QY 241 ATGCTGGAGACTGGAGAAATACAAAGACGACGAAAGCTCATAGTCGAGCGTACAG 300
Db 403 ATGCTGGAGACTGGAGAAATACAAAGACGACGAAAGCTCATAGTCGAGCGTACAG 462
QY 301 GGAATGCCCATGAACATCCGGGGCCCATGTGTCTAGTCTCTCTGAACATTGAGGAATG 360
Db 463 GGAATGCCCATGAACATCCGGGGCCCATGTGTCTAGTCTCTCTGAACACTGAGGAATG 522
QY 361 AAGTGAAAAACCCCGAAGATACCATCATGAAGGAGAAAGGGCAAGAGTCACTGAG 420
Db 523 AAGTGAAAAACCCCGAAGATACCATCATGAAGGAGAAAGGGCAAGAGTCACTGAG 582
QY 421 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTTAAGAACATATATTCTTCAG 480
Db 583 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTTAAGAACATATATTCTTCAG 642
QY 481 GATGATACGGAACAAGCAGCGGGAATCTCTCACATCTCTCTGATATGAGAGATAT 540
Db 643 GATGATACGGAACAAGCAGCGGGAATCTCTCACATCTCTCTGATATGAGAGATAT 702
QY 541 AACCAGAGTGGGCTACTGACAGGACTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 703 AACCAGAGTGGGCTACTGACAGGACTGAGCCACATCGCCGCTTGTCTCTCTAT 762
QY 601 CTTCTGAGGAGATGATCTTGGGCACTGTGACGCTGTGCGCAGTGAGAGGCACTTC 660
Db 763 CTTCTGAGGAGATGATCTTGGGCACTGTGACGCTGTGCGCAGTGAGAGGCACTTC 822
QY 661 CTGACGGAATTTACAGCCCAAAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
Db 823 CTGACGGAATTTACAGCCCAAAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 882
QY 721 CATGTGTAGCCACGTCAACCAAGCAAGCATGGGCAATGAGGCAAGAAAGATCTATGT 780
Db 883 CATGTGTAGCCACGTCAACCAAGCAAGCATGGGCAATGAGGCAATGAGGCAATG 923
QY 781 GGGCAGTGTCCCGGTAGGCTGCTCATCCGATATTGATGACGGATCTCTCGGG 840
Db 924 -----GATCTCTCTCGGG 936
QY 841 CTCACCTTGGGCTGTGGGACGTGTATCTGTAGAAAGGCGAAGCGGTGATGCGGATA 900
Db 937 CTCACCTTGGGCTGTGGGACGTGTATCTGTAGAAAGGCGAAGCGGTGATGCGGATA 996
QY 901 ACAAGATCGCTTTAAGGTTACAGAGAAGCGCTTACGAAGACGTCCAGGTGTGCGCG 960
Db 997 ACAAGATCGCTTTAAGGTTACAGAGAAGCGCTTACGAAGACGTCCAGGTGTGCGCG 1056
QY 961 TGGCAGCTTTTGCACACCGGTTCTGTATACCTGGGCCAGGGATGAGACACTGTGCTC 1020
Db 1057 TGGCAGCTTTTGCACACCGGTTCTGTATACCTGGGCCAGGGATGAGACACTGTGCTC 1116
QY 1021 AAGCATCTTAGGGCTCTATGAAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1117 AAGCATCTTAGGGCTCTATGAAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1176
QY 1081 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCGGCTTCACTGTGCGGAAAG 1140
Db 1177 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCGGCTTCACTGTGCGGAAAG 1236
QY 1141 ACCCTTGCAGAGGGGACAGGCGGCCCCCTCCAGGCCCAAGCCCGGTTCCCGGGCCC 1200
Db 1237 ACCCTTGCAGAGGGGACAGGCGGCCCCCTCCAGGCCCAAGCCCGGTTCCCGGGCCC 1296
QY 1201 ATTGTGACCTTCCCGCCAGCGGCACTCGTTCTTCCACACCTGTCTGTGGGGCT 1260
Db 1297 ATTGTGACCTTCCCGCCAGCGGCACTCGTTCTTCCACACCTGTCTGTGGGGCT 1356
QY 1261 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
Db 1357 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1416
QY 1321 GGAGACCTCAGGGTCTCTGAGATTCTGCAGTGGAATCCATGCCCCGCTCCCAAG 1380

Db 1417 GGAGGACCTCAGGGTTCCTGAGATTCCTGCACTGGAATCCATGCCCGCCTCCCAACG 1476
QY 1381 GACCTGACGTAGAGGGCCCTTGTTCCGCATTATGATTTTCAGACAGAGCTGCTGGTTC 1440
Db 1477 GACCTGACGTAGAGGGCCCTTGTTCCGCATTATGATTTTCAGACAGAGCTGCTGGTTC 1536
QY 1441 CGTGCCATATCCCAAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCTGGAG 1500
Db 1537 CGTGCCATATCCCAAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCTGGAG 1596
QY 1501 CGGGTAGATCGGCTTTTCGCTGCACCCCAAGCATGATTCGACAGGGCACCCCTTCAGA 1560
Db 1597 CGGGTAGATCGGCTTTTCGCTGCACCCCAAGCATGATTCGACAGGGCACCCCTTCAGA 1656
QY 1561 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGCGCTTGCTGCGGCTCCACTTG 1620
Db 1657 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGCGCTTGCTGCGGCTCCACTTG 1716
QY 1621 GAAAGTTCAGTCCCTCCAGGCTTCTAGAAGCATCTGGGCGAGGGCTCATGCTGAT 1680
Db 1717 GAAAGTTCAGTCCCTCCAGGCTTCTAGAAGCATCTGGGCGAGGGCTCATGCTGAT 1776
QY 1681 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCGTCTCGTTTATTTTGGTTAA 1740
Db 1777 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCGTCTCGTTTATTTTGGTTAA 1836
QY 1741 CTTATGAAAAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
Db 1837 CTTATGAAAAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1896
QY 1801 AGACCCCAAGATCAAAAGCCAAACCATGCCAGCCCCCTCCAGCACCCCCCAGCCAGCAC 1860
Db 1897 AGACCCCAAGATCAAAAGCCAAACCATGCCAGCCCCCTCCAGCACCCCCCAGCCAGCAC 1956
QY 1861 CATCGTCTGAATCTGACGACACCGTGAGCCTGCTTGTACTTTAAACTCATGAAG 1920
Db 1957 CATCGTCTGAATCTGACGACACCGTGAGCCTGCTTGTACTTTAAACTCATGAAG 2016
QY 1921 ATAACCTACCTCAAGTTTGAATAAATGTTCTGTTGAATG 1964
Db 2017 ATAACCACTTCATGTTTGAATAAATGTTCTGTTGAATG 2060

RESULT 5
US-10-108-260A-676
; Sequence 676, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-676

Query Match 92.0%; Score 1807.4; DB 6; Length 2647;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
QY 1 ATGACGTGTAGAGTCCGGGCAAGTGTGTGGGCACAAGAGCGAGAGACATCATTAATG 60
Db 755 ATGACGTGTAGAGTCCGGGCAAGTGTGTGGGCACAAGAGCGAGAGACATCATTAATG 814
QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTAAGCCTTTTCGA 120
Db 815 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTAAGCCTTTTCGA 874

QY 121 AGCTAACAAACAACGTCGATCAATTGGGGATTGTACATGAGACGAGCTGCCCTCTCTG 180
Db 875 AGCTAACAAACAACGTCGATCAATTGGGGATTGTACATGAGACGAGCTGCCCTCTCTG 934
QY 181 ACTGCGCGGAGGCGGAAGCAAAATCCGGCGGAGATCAGCCGAAAGACAAGTGGTGGAT 240
Db 935 ACTGCGCGGAGGCGGAAGCAAAATCCGGCGGAGATCAGCCGAAAGACAAGTGGTGGAT 994
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGACAGAAAGCTCATAGATCGAGCGTACAAG 300
Db 995 ATGCTGGAGACTGGGAGAAATACAAAAGCAGACAGAAAGCTCATAGATCGAGCGTACAAG 1054
QY 301 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAAATG 360
Db 1055 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAAATG 1114
QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCATCTGAG 420
Db 1115 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCATCTGAG 1174
QY 421 CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAGAAAGCATATATCTTCAGG 480
Db 1175 CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAGAAAGCATATATCTTCAGG 1234
QY 481 GATGATACGGAACCAAGCAGCGCGGAATCTCCACATCTCTCTGCGATATGAGAGTAT 540
Db 1235 GATGATACGGAACCAAGCAGCGCGGAATCTCTCCACATCTCTCTGCGATATGAGAGTAT 1294
QY 541 AACCAGAGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 1295 AACCAGAGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 1354
QY 601 CTTCCTGAGAGGATGATCTGCGGCACTGTGCACTGTGCGCAGTGAGAGGCACTCC 660
Db 1355 CTTCCTGAGAGGATGATCTGCGGCACTGTGCACTGTGCGCAGTGAGAGGCACTCC 1414
QY 661 CTGAGGGATTTTCACAGCCCAAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 720
Db 1415 CTGAGGGATTTTCACAGCCCAAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 1474
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATGT 780
Db 1475 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCA----- 1515
QY 781 GGGCAGTGTCCCCGTTAGGCTGCCCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
Db 1516 -----GATCTCTCTCGGG 1528
QY 841 CTCACCTGCGCCTGTGGAGCGTGTATCTGTAGAAAGCGAAACAGGCGTTGATGCCGATA 900
Db 1529 CTCACCTGCGCCTGTGGAGCGTGTATCTGTAGAAAGCGAAACAGGCGTTGATGCCGATA 1588
QY 901 ACAAGATCGCCTTTAAGGTTCAAGCAGAAAGCGCTCAAGAAAGCTCCAGGTGTGCCCG 960
Db 1589 ACAAGATCGCCTTTAAGGTTCAAGCAGAAAGCGCTCAAGAAAGCTCCAGGTGTGCCCG 1648
QY 961 TGGCAGCTTTTGGCAACCGGTCGTTGATACCTGGGCCAGGGATGAGAGCACTGTGCTC 1020
Db 1649 TGGCAGCTTTTGGCAACCGGTCGTTGATACCTGGGCCAGGGATGAGAGCACTGTGCTC 1708
QY 1021 AAGCATTTAAGGCGCTTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCACCCCCA 1080
Db 1709 AAGCATTTAAGGCGCTTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCACCCCCA 1768
QY 1081 GCCAAACCGAGCAAGGTCGTGGCATCCAGGCGCTGTGCCGCTTCACTGGCGGGAAG 1140
Db 1769 GCCAAACCGAGCAAGGTCGTGGCATCCAGGCGCTGTGCCGCTTCACTGGCGGGAAG 1828
QY 1141 ACCCTTGCAAGGGGACAGGAGCGCCCTCCAGGCCCAACAGCCCGTTCCGCGGGCCC 1200
Db 1829 ACCCTTGCAAGGGGACAGGAGCGCCCTCCAGGCCCAACAGCCCGTTCCGCGGGCCC 1888

QY	1201	ATTGGTCAGCTTCCCCGCCACGGGGCACTCGTTCTTCCACACCCTGTCTGTGGGCT	1260
Db	1889	ATTGGTCAGCTTCCCCGCCACGGGGCACTCGTTCTTCCACACCCTGTCTGTGGGCT	1948
QY	1261	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG	1320
Db	1949	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG	2008
QY	1321	GGAGGACCTCAGGGTTCGTGAGATTCTCGAGTGAATCCATGCCCCGCTCCACG	1380
Db	2009	GGAGGACCTCAGGGTTCGTGAGATTCTCGAGTGAATCCATGCCCCGCTCCACG	2068
QY	1381	GACCTGGACGTAGAGGGCCCTTGTTGCTCCGCCATTATGATTTCAACAAGCTGCTGGTTC	1440
Db	2069	GACCTGGACGTAGAGGGCCCTTGTTGCTCCGCCATTATGATTTCAACAAGCTGCTGGTTC	2128
QY	1441	CGTGCCATATCCCAAGAGACCAAGCTGGCCCCCTGCTGGCAGGGCTGAACACCCCTGCGAG	1500
Db	2129	CGTGCCATATCCCAAGAGACCAAGCTGGCCCCCTGCTGGCAGGGCTGAACACCCCTGCGAG	2188
QY	1501	CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA	1560
Db	2189	CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA	2248
QY	1561	GCTAGGGACGAACAGCCGCTGTGCTCCACCTCAGGGCCCTTGCCCTTGCGGCTCCACTTG	1620
Db	2249	GCTAGGGACGAACAGCCGCTGTGCTCCACCTCAGGGCCCTTGCCCTTGCGGCTCCACTTG	2308
QY	1621	GAAAGTTCTCAGTTCCTCCACAGGCTTCTAGAAGCATTTGGGCCAAGGCTCATGGCTGAT	1680
Db	2309	GAAAGTTCTCAGTTCCTCCACAGGCTTCTAGAAGCATTTGGGCCAAGGCTCATGGCTGAT	2368
QY	1681	AATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTGCCTGCTCTGTTATTTTGGTTAAA	1740
Db	2369	AATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTGCCTGCTCTGTTATTTTGGTTAAA	2428
QY	1741	CTTATGAAATGTATTAAGAAAGAGTGCAGCTCGAAGAGATTAGAGATGAACACACC	1800
Db	2429	CTTATGAAATGTATTAAGAAAGAGTGCAGCTCGAAGAGATTAGAGATGAACACACC	2488
QY	1801	AGACCCCAAGATCACAAGCCAACCATGCCCCAGCCCCCTCCAGCACCCCCAAGCCCAAGAC	1860
Db	2489	AGACCCCAAGATCACAAGCCAACCATGCCCCAGCCCCCTCCAGCACCCCCAAGCCCAAGAC	2548
QY	1861	CATCGTCTGAATTTCTGACGACACCCGTGAGCCCTGCTTTGTACTTTAACTCATGGAAG	1920
Db	2549	CATCGTCTGAATTTCTGACGACACCCGTGAGCCCTGCTTTGTACTTTAACTCATGGAAG	2608
QY	1921	ATAACTACCTTCACGTTTTGAATAAATGTTTCCCTGTTG	1959
Db	2609	ATAACCACCTTCATGTTTTGAATAAATGTTTCCCTGTTG	2647

```

RESULT 6
US-10-071-838-3
; Sequence 3, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularex Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071, 838
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267, 615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2146

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1832)
;
US-10-071-838-3

```

Query Match	91.3%	Score 1794;	DB 5;	Length 2146;
Best Local Similarity	99.9%	Pred. No. 0;		
Matches 185;	Conservative	0;	Mismatches 0;	Indels 1;
				Gaps 1;

QY	159	TGAGACGGAGCTGCCTCTCTCTGACTGCGCGGGAGGCGAAGCAAAATTCGGCGGGAGATCAG	218
Db	342	TGAGACGGAGCTGCCTCTCTCTGACTGCGCGGGAGGCGAAGCAAAATTCGGCGGGAGATCAG	401
QY	219	CCGAAAGAGCAAGTGGGTGATATGCTGGGAGACTGGAGAAATACAAAGCAGCAGAAA	278
Db	402	CCGAAAGAGCAAGTGGGT - GATATGCTGGGAGACTGGAGAAATACAAAGCAGCAGAAA	460
QY	279	GCTCATAGATCGAGCGTACAAAGGAATGCCATGAACATCCGGGGCCCGATGTGTCAGT	338
Db	461	GCTCATAGATCGAGCGTACAAAGGAATGCCATGAACATCCGGGGCCCGATGTGTCAGT	520
QY	339	CCTCCTGAACATTGAGGAAATGAAGTTGAAAAACCCCGGAAGATACCAGATCATGAAGGA	398
Db	521	CCTCCTGAACATTGAGGAAATGAAGTTGAAAAACCCCGGAAGATACCAGATCATGAAGGA	580
QY	399	GAAAGGCAAGAGGTCACTGAGCACATCCAGCGCATCGACCGGGACGTAAAGCGGGACATT	458
Db	581	GAAAGGCAAGAGGTCACTGAGCACATCCAGCGCATCGACCGGGACGTAAAGCGGGACATT	640
QY	459	AAGGAAGCATATATTCTTCAGGGATCGATACGGAACCAAGCAGCGGAACTACTCCACAT	518
Db	641	AAGGAAGCATATATTCTTCAGGGATCGATACGGAACCAAGCAGCGGAACTACTCCACAT	700
QY	519	CCTCCTGCATATGAGAGATATAACCCGAGGTGGCTACTGCAGGGACCTGAGCCACAT	578
Db	701	CCTCCTGCATATGAGAGATATAACCCGAGGTGGCTACTGCAGGGACCTGAGCCACAT	760
QY	579	CGCCGCTTGTCTCTCTATCTTCTGAGGAGATGCATTCTGGGCACTGTGCAGCT	638
Db	761	CGCCGCTTGTCTCTCTATCTTCTGAGGAGATGCATTCTGGGCACTGTGCAGCT	820
QY	639	GCTGGCCAAGTAGAGGCACTCCCTGCAGGGATTTCACAGCCCAATGGCGGGACCGTCCA	698
Db	821	GCTGGCCAAGTAGAGGCACTCCCTGCAGGGATTTCACAGCCCAATGGCGGGACCGTCCA	880
QY	699	GGGGCTCCAAGACCAACAGGAGCATGTGTAGCCACGTCACAACCCAAAGCATTGGGGCA	758
Db	881	GGGGCTCCAAGACCAACAGGAGCATGTGTAGCCACGTCACAACCCAAAGCATTGGGGCA	940
QY	759	TCAGGACAAAGAAAGATCTATGTGGGCAGTGTCCCGTTAAGCTGCTCATCCGATATT	818
Db	941	TCAGGACAAAGAAAGATCTATGTGGGCAGTGTCCCGTTAAGCTGCTCATCCGATATT	1000
QY	819	GATTGACGGATCTCTCGGGCTCACCCCTGCGCTGTGGACGTGTATCTGTAGAAG	878
Db	1001	GATTGACGGATCTCTCGGGCTCACCCCTGCGCTGTGGACGTGTATCTGTAGAAG	1060
QY	879	CGAACAGGCGTTGATGCCGATACACAGAATCGCCTTTAAGTTACAGCAGAGCGCCTCAC	938
Db	1061	CGAACAGGCGTTGATGCCGATACACAGAATCGCCTTTAAGTTACAGCAGAGCGCCTCAC	1120
QY	939	GAAAGCCTCCAGGTGTGGCCCGTGGGCACGTTTTTGCAACCCGTTGTTGATACCTGGGC	998
Db	1121	GAAAGCCTCCAGGTGTGGCCCGTGGGCACGTTTTTGCAACCCGTTGTTGATACCTGGGC	1180
QY	999	CAGGGATGAGGACATGTGCTCAAGCATCTTAAGGCTCTATGAAGAACTAACAGAAA	1058
Db	1181	CAGGGATGAGGACATGTGCTCAAGCATCTTAAGGCTCTATGAAGAACTAACAGAAA	1240

QY 721 CATGTGTAGCCAGCTCAACAACCCAAAGACATGGGGCATCAGACGAAGAAGATCTATGT 780
Db 721 CATGTGTAGCCAGCTCAACAACCCAAAGACATGGGGCATCA-----761
QY 781 GGGCAGTGTCCCGGTAGGCTGCTCATCCGGATATTGATTGACGGGATCTCTCGGG 840
Db 762 -----761
QY 841 CTCACCCCTGGCCTGTGGGACGTATCTGTAGAAGGCGAAGCGCTTGATGCCGATA 900
Db 762 -----GTATCTGTAGAAGGCGAAGCGCTTGATGCCGATA 798
QY 901 ACAAGATCGCCTTTAAGGTTCAAGCAGAAGCGCCTCACGAAGAGCTCCAGTGTGGCCCG 960
Db 799 ACAAGATCGCCTTTAAGGTTCAAGCAGAAGCGCCTCACGAAGAGCTCCAGTGTGGCCCG 858
QY 961 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC 1020
Db 859 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC 918
QY 1021 AAGCATCTTAGGCTCTATAGAAAGAACTAACAAAGAACAGGGGACCTGCAACCCCA 1080
Db 919 AAGCATCTTAGGCTCTATGAAAGAACTAACAAAGAACAGGGGACCTGCAACCCCA 978
QY 1081 GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCGGCTTCACTGCGGGAAG 1140
Db 979 GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCGGCTTCACTGCGGGAAG 1038
QY 1141 ACCCTCTGCAAGGGGGACAGGCAAGGCCCCCTCCAGGCCACCAAGCCGGTTCCCGCGGCC 1200
Db 1039 ACCCTCTGCAAGGGGGACAGGCAAGGCCCCCTCCAGGCCACCAAGCCGGTTCCCGCGGCC 1098
QY 1201 ATTTGTGAGCTTCCCCCGCACCGGCACTGTTCTTCCACACCTGTCTGTGGGCT 1260
Db 1099 ATTTGTGAGCTTCCCCCGCACCGGCACTGTTCTTCCACACCTGTCTGTGGGCT 1158
QY 1261 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG 1320
Db 1159 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG 1218
QY 1321 GGAGACCTCAGGGTCTCTGAGATTCTGCAAGTGAACCTGACCCCGCTCCCAAG 1380
Db 1219 GGAGACCTCAGGGTCTCTGAGATTCTGCAAGTGAACCTGACCCCGCTCCCAAG 1278
QY 1381 GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTTCAACAGAGCTGTGGGTC 1440
Db 1279 GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTTCAACAGAGCTGTGGGTC 1338
QY 1441 CGTGCCATATCCAGAGAGACAGCTGAGCCCCCTGTGGCAGGCTGAACACCCCTGGGAG 1500
Db 1339 CGTGCCATATCCAGAGAGACAGCTGAGCCCCCTGTGGCAGGCTGAACACCCCTGGGAG 1398
QY 1501 CGGTTGAGATCGGCTTTGCTGACACCCAGCACTGATTCCGACAGGGCAACCCCTTCAGA 1560
Db 1399 CGGTTGAGATCGGCTTTGCTGCTGACACCCAGCACTGATTCCGACAGGGCAACCCCTTCAGA 1458
QY 1561 GCTAGGAGCAAGCAGCGTGTGCTCCCACTCAGGGCCTTGCTTGCGGCTCACTTG 1620
Db 1459 GCTAGGAGCAAGCAGCGTGTGCTCCCACTCAGGGCCTTGCTTGCGGCTCACTTG 1518
QY 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
Db 1519 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1578
QY 1681 AATTTCCCTAGGCTTAACAACCCCAAGAGCTTGCGCTCTGTTTATTGTTGTTAAA 1740
Db 1579 AATTTCCCTAGGCTTAACAACCCCAAGAGCTTGCGCTCTGTTTATTGTTGTTAAA 1638
QY 1741 CTTATGAAAAATGTATTAAGAAAGATGAGCTCGAGAGAGATTCAAGATGAACACACC 1800
Db 1639 CTTATGAAAAATGTATTAAGAAAGATGAGCTCGAGAGAGATTCAAGATGAACACACC 1698
QY 1801 AGACCCCAAGATCAAAAGCCAACGATGCCAGCCCTCCAGCAACCCCAAGCCCAAGAC 1860

Db 1699 AGACCCCAAGATCAAAAGCCAACCATGCCAGCCCTCCAGACACCCCAAGCCCAAGAC 1758
QY 1861 CATGTTCTGAATTCTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG 1920
Db 1759 CATGTTCTGAATTCTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG 1818
QY 1921 ATAACCTTCAAGCTTTTGAATAATGTTTCTGTTGAAATG 1964
Db 1819 ATAACCTTCAAGCTTTTGAATAATGTTTCTGTTGAAATG 1862
RESULT 8
US-10-094-466-61
; Sequence 61, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spytek et al.
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 61
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31)..(1678)
US-10-094-466-61
Query Match 83.4%; Score 1638.8; DB 6; Length 1752;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1670; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1 ATGACGTGTGTAAGGTGCGCGGCACTGTGGGCAACAAGCGAGACATCATATATG 60
Db 31 ATGACGTGTGTAAGGTGCGCGGTAATTGTTGGGCAACAAGCGAGACATCATATATG 90
QY 61 AATATCGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 120
Db 91 AATATCGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 150
QY 121 AGCTACAACAACGTCATGATTTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
Db 151 AGCTACAACAACGTCATGATTTTGGGATTGTACAGAGACGAGCTGCTCTCTG 210
QY 181 ACTGCGGGGAGGGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240

Db 211 ACTGCGGAGAGTGAAGCAATTCCGCGGAGATCAGCCGAAAGCAAGTGGTGA 270
Qy 241 ATGCTGGAGACTGGAGAAATACAAGACAGAAAGCTCATATGAGCGTACAAG 300
Db 271 ATGCTGGAGAAATGGAGACACTACAAGAAACAGAAAGCTCATATGAGCGTACCAG 330
Qy 301 GGAATGCCATGAACATCCGGGGCCCCGATGTGTCAGTCTCTGAAACATTGAGAAATG 360
Db 331 GGAATTCCTATGAACATCCGGGGCCCCGATGTGTCAGTCTCTGAAACATTGAGAAATC 390
Qy 361 AAGTGAAGAAACCCCGAAGATACCATCATGAAGAGAGAGGAGAGTCACTGAG 420
Db 391 AAGTGAAGAAACCCCGAAGATACCATCATGAAGAGAGGAGAGTCACTGAA 450
Qy 421 CACATCCAGCGCATCGACCGGAGCGTAAGCGGACATTAAAGAGCATATATTCTTCAG 480
Db 451 CACATCCAGCAGATGGACCTGAGCGTAAGCGGACATTAAAGAGCATATATTCTTCAG 510
Qy 481 GATCGATACGGAACCAAGCAGCGGGAACCTCAACATCTCTGCGCATATGAGAGTAT 540
Db 511 GATCGATACGGAACCAAGCAGCGGGAACCTTAACATCTCTCTGCGCATATGAGAGTAT 570
Qy 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 571 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 630
Qy 601 CTTCCTGAGAGGATGATCTCTGGGCACTGTGTCAGCTGTGCGCAGTGAGAGCACTCC 660
Db 631 CTTCCTGAGAGGATGATCTCTGGGCACTGTGTCAGCTGTGCGCAGTGAGAGCACTCC 690
Qy 661 CTGCAAGGATTTCAAGCCCAATGCGGAGCGTCCAGGGCTCAAGACCAAGAGAG 720
Db 691 CTGCAAGGATTTCAAGCCCAATGCGGAGCGTCCAGGGCTCAAGACCAAGAGAGAG 750
Qy 721 CATGTGTAGCCAGCTCAACCAAGACCATGGGCACTCAGACAAGAAAGATCTATGT 780
Db 751 CATGTGTAGCCAGCTCACTACCAACACCATGTGGCATCAGACAAGAAAGATCTATGT 810
Qy 781 GGGCAGTGTCCCCCTTAGGCTGCTCATCCGGATATTGATTGACCGGATCTCTCGGG 840
Db 811 GGGCAGTGTTCGTCCTTAGGCTGCTCATCCGATATTGATTGACCGGATCTCTCTCGGG 870
Qy 841 CTCACCCCTGCGCTGTGGGACGTGTATCTGGTAGAAGGCGAAGCGGTTGATGCCGATA 900
Db 871 CTCACCCCTGCGCTGTGGGACGTGTATCTGGTAGAAGGCGAAGCGGTTGATGCCGATA 930
Qy 901 ACAAGATCGCCTTTAAGGTTCAAGCAGAAAGCGCTCAAGAAAGCTCAGGTGTGCCCG 960
Db 931 ACAAGATCGCCTTTAAGGTTCAAGCAGAAAGCGCTCAGAAAGCTCAGGTGTGCCCG 990
Qy 961 TGGGCAAGTTTTCGAACCGGTTGTTGATTAACCTGGGCGAGGATGAGGACACTGTGCTC 1020
Db 991 TGGGCAAGTTTTCGAACCGGTTGTTGATTAACCTGGGCGAGGATGAGGACACTGTGCTC 1050
Qy 1021 AAGCATCTTAGGGCTCTATGAAGAAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1051 AAGCATCTTAGGGCTCTATGAAGAAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1110
Qy 1081 GCCAAACCCGAGCAAGGGTCTGCGGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG 1140
Db 1111 GCCAAACCCGAGCAAGGGTCTGCGGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG 1170
Qy 1141 ACCCTCTGCAAGGGGAGCAGGAGGCCCTTCAGGCCACACCGGTTCCCGCGGCC 1200
Db 1171 ACCCTCTGCAAGGGGAGCAGGAGGCCCTTCAGGCCACACCGGTTCCCATGGCC 1230
Qy 1201 ATTGCTAGCTTCCCGCAGCGGCACTGTCTTCCACACCTGTCTGTGGGCT 1260
Db 1231 ATTGCTAGCTTCCCGCAGCGGCACTGTCTTCCACACCTGTCTGTGGGCT 1290
Qy 1261 GTCCGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCCCTGTGCTAG 1320

Db 1291 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCCCTGGCTCAG 1350
Qy 1321 GGAGGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCTCATGCCCCGCTCCCAAG 1380
Db 1351 GGAGGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCTCATGCCCCGCTCCCAAG 1410
Qy 1381 GACCTGACCTAGAGGGCCCTTGGTTCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1440
Db 1411 GACCTGACCTAGAGGGACCTTGGTTCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1470
Qy 1441 CGTGCAATCCAGGAGGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCTGGAG 1500
Db 1471 CGTGCAATCCAGGAGGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCCTGGAG 1530
Qy 1501 CGGTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1560
Db 1531 CGGTGAGATCGGCTTTCAGTCCGCCCAAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1590
Qy 1561 GCTAGGAGCAACAGCCGCTGTGCTCCCACTCAGGGCCTTGGCTTGCAGCTTG 1620
Db 1591 GCTAGGAGCAACAGCAGTGTGCTCCCACTCAGGACCTTGTGCGGCTTCACCTTG 1650
Qy 1621 GAAAGTCTCAGTTCCTCCAGGCTTCTAGAGCATGTGGCCAGGCTCATGGCTGAT 1680
Db 1651 GAAAGTCTCAGTTCCTCCAGGCTTCTAGAGCATGTGGCCAGGCTCATGGCTGAT 1710
Qy 1681 AATTCCCTAGGCTTAAACCCCAAGCAAGCTTGCCTCTC 1722
Db 1711 AATTCCCTAGGCTTAAACCCCAAGCAAGCTTGCCTCTC 1752

RESULT 9
US-10-450-763-21335
; Sequence 21335, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 21335
; LENGTH: 7856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1697)..(4054)
; OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number
; OTHER INFORMATION: X63546, Smith-Waterman Score=4233.
US-10-450-763-21335
Query Match 62.1%; Score 1220; DB 9; Length 7856;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
Qy 1 ATGACGTGTAGAGGTGCGCGGCAAGTGTGGGCAAGAGGAGAGACATCATATG 60
Db 1697 ATGACATGTGTAGAGATGACAGTATGTTGACGACACAGAGCGGAAGACATATATG 1756
Qy 61 AAATACGAAGAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGGCTAAGCTTTTGA 120
Db 1757 AAGTATGACAAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGGCTGAAGCCGTTGA 1816
Qy 121 AGCTACAACAACAGCTGATCATTTGGGATTGTATCATGAGACGAGCTGCTCTCTG 180

```
Db 1817 ATCAACAGCAGCA---TTGATCGTTTGGCATTTTGCATGAGACGGAGCTGCCTCCTGTG 1873
QY 181 ACTGCGGGGAGGCGCAAGCAATTCGCGGGGAGATCAGCCGAAGCAAGTGGGTGAT 240
Db 1874 ACTGACGGGAGGCGAAGAAAATTCGCGGGAGATGACACGAACGACAAAGTGATGAA 1933
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAAAAGCTCATAGATCGAGCTACAG 300
Db 1934 ATGCTGGAGAAATGGAGACATATAGCAGTAGCAAACTCATAGATCGAGTGTACAG 1993
QY 301 GGAATGCCCATGAACATCCGGGGCCGATGTGCTCAGTCCCTCTGAACATTGAGAAATG 360
Db 1994 GGAATTCCTCATGAACATCCGGGGCCGAGTGTGCTCAGTCCCTCTGAACATTGAGAAATC 2053
QY 361 AAGTGAACAAACCCCGAAGATACCAAGATCATGAAGAGAGAGGCAAGAGTCACTGAG 420
Db 2054 AAGTGAACAAACCCCGAAGATACCAAGATCATGAAGAGAGGCGCAAGAGTCACTGAA 2113
QY 421 CACATCCAGCGCATCGACCCGGGACGTAGCGGGACATTAAGAAACATATATTCTTCAG 480
Db 2114 CACATCCACCATTCGACCTTGACGTGAGACGACTCTCCGAAACATGTCTTTAGG 2173
QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTCCACATCCTCTGCAATATGAGAGAT 540
Db 2174 GATCGATATGAGCCAAAGCAGAGGGAATCTACATCCTCTGCTTATTCGAGATAT 2233
QY 541 AACCCGAGGTGGGCTACTGACGGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 600
Db 2234 AACCCGAGGTGGGCTACTGACGGGACCTGAGCCACATCACCGCTGTCTCTCTAT 2293
QY 601 CTTCTGAGAGGATGATCTGCGGCACTGTGACGTGCTGCGCAGTGAGAGGCACTCC 660
Db 2294 CTGCTGAGAGGAGCAGCATTCGCGCACTGTGACGTGCTGCGCAGTGAGAGGCACTCC 2353
QY 661 CTGCAAGGATTTCAAGCCCAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
Db 2354 CTGCCAGGATTCACAGCCCAATGCTGGGACAGTCCAGGGGCTCCAAGACCAAGAG 2413
QY 721 CATGTGTAGCCACGTCAACACCCCAAGCCATGGGGCATCAGGCAAGAAAGATCTATGT 780
Db 2414 CATGTGTACCCAAGTCAACACCCCAAGCCATGTGGCATCAGGCAAGAAAGTCTATGC 2473
QY 781 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGATATGTATGACGGGATCTCTCGGG 840
Db 2474 GGGCAGTGTCCCGCTTAGGCTGCTCTCCGGAACCTGATGACGGGATCTCTCGGG 2533
QY 841 CTCACCCCTGCGCTGTGGAGCGTGTCTGTAGAAAGGCAACAGCGTTGATGCCGATA 900
Db 2534 CTCACCCCTGCGCTGTGGAGCGTGTATTGTGGAAGAGAACAGGTGTGATGCCAATA 2593
QY 901 ACAAGAAATCGCTTTAAGGTTCAAGCAGAGCGCTCAAGAACAGTCCAGGTGTGCGCCG 960
Db 2594 ACCAGCATGTCTTTAAGGTTCAAGCAGAGCGCTCATGAAGACATCCAGGTGTGCGCTG 2653
QY 961 TGGGCACTTTTTCGAACCGGTTCTGTGATACCTGGGCAAGGATGAGACACTGTGCTC 1020
Db 2654 TGGGCACTGTGCGGAACCAATTTCTGATACCTGGGCAATGAACGATGACACCGTGTCTC 2713
QY 1021 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 1080
Db 2714 AAGCATCTTAGGGCTCTACGAAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 2773
QY 1081 GCCAAACCGGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCAGTGGCGGGAAG 1140
Db 2774 GCCAAACCGGAGCAAGGTCCTTGGCACCCAGGCTGTGCGGCTTCAGTGGTGGGAAG 2833
QY 1141 ACCCTTGGAAGGGGAGCAGCAGGCGCTCCAGGCCACAGCCCGTTCCCGCGGCC 1200
Db 2834 ACCCTTGGAAGGGGTATAGCAGGCGCTCCAGGCCACAGCCCAATTCAGCGCGGCC 2893
QY 1201 ATTTGTGACCTTCCCGCCACGGGCACTGTCTTCCACACCCTGTCTGTGGGGGCT 1260
|||||
```

```
Db 2894 ATTTGCTAGCTTCCCGCCATGSGGCACTCTGTTTTTCCAGCGCCTGTCTGTGGGCT 2953
QY 1261 GTCCGGGAAGACCTAACCTGTGGGCACTCAGGGGTGTGCCAGCCCGCCCTGCTCAG 1320
Db 2954 GTCCGGGAAGACGTAACCTGTGGGCACTCAGGGGTGTGCCAGCCCTGCGCTCAG 3013
QY 1321 GGAGACCTCAGGGTCTCTGAGATTCCTGCACTGGAACCTCCATGCCCGCTCCCAAG 1380
Db 3014 GGAGACCTCAGGGTCTCTGAGATTCCTGAGATTCCTGAGTGAAGTCAATGCCCGCTCCCAAG 3073
QY 1381 GACCTGACGTAGAGGGCCCTTGTGTTCCGCCATTATGATTTGAGACAGAGCTGTGGTC 1440
Db 3074 GACCTGATATAGGGGCGCTTGTGTTCCGCCATTATGATTTGAAACGAGCTGTGGTC 3133
QY 1441 CGTGCCATATCCAGAGGAGACCAAGCTGCGCCCTGCTGCGAGCTGAACACCTGCGGAG 1500
Db 3134 CGTGCCATATCCAGAGGAGACCAAGCTGCGCCACCTGCTGCGAGGCTGAACACTGCGGAGAG 3193
```

```
RESULT 10
US-10-450-763-24017
; Sequence 24017, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 24017
; LENGTH: 7856
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (1697)..(4054)
; OTHER INFORMATION: 100% homologous to Homo sapiens oncogene, accession number
; US-10-450-763-24017
```

```
Query Match 62.1%; Score 1220; DB 9; Length 7856;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 1 ATGACGTGTAGAGGTGCGGGCAGTTGTGGGCAAGAGCGAGAGACATCATATG 60
Db 1697 ATGACATGTGTAGAGATGCAATAGTTTGACGGCACAGAGCGGAAGACATCTTATG 1756
QY 61 AAATACGAAAAGGACACCCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTCA 120
Db 1757 AAGTATGACAAAGGACACCCGAGCTGGGCTGCCAGAGACAAAGGGCCTGAAGCCGTTTGA 1816
QY 121 AGCTACAACAACAGTCATCTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
Db 1817 ATCAACAGCAGCA--TTGATCGTTTGGCATTTTGCATGAGACGAGCTGCTCTGTG 1873
QY 181 ACTGCGGGGAGGCGAAGCAATTCGCGGAGATCAGCCGAAGCAAGTGGGTGAT 240
Db 1874 ACTGACGGGAGGCGAAGAAATTCGCGGAGATGACAGAAAGCAAGTGGATGAA 1933
QY 241 ATGCTGGAGACTGGAGAAATACAAAAGCAGCAAGAAAGCTCATAGATCGAGCTGACAG 300
Db 1934 ATGCTGGAGAAATGGAGACATATAGCAGTAGCAAACTCATAGATCGAGTGTACAG 1993
QY 301 GGAATGCCCATGAACATCCGGGGCCGATGTGCTCAGTCCCTCTGAACATTGAGAAATG 360
|||||
```


Db	1994	GGAATTC	2053
Qy	361	AAAGTGA	420
Db	2054	AAAGTGA	2113
Qy	421	CACATCC	480
Db	2114	CACATCC	2173
Qy	481	GATGATA	540
Db	2174	GATGATA	2233
Qy	541	AACCCGA	600
Db	2234	AACCCGA	2293
Qy	601	CTTCCTG	660
Db	2294	CTGCCTG	2353
Qy	661	CTGCAGG	720
Db	2354	CTGCAGG	2413
Qy	721	CATGTGT	780
Db	2414	CATGTGT	2473
Qy	781	GGGAGTG	840
Db	2474	GGGAGTG	2533
Qy	841	CTCACCT	900
Db	2534	CTCACCT	2593
Qy	901	ACAAGAT	960
Db	2594	ACCAAGC	2653
Qy	961	TGGGACG	1020
Db	2654	TGGGACG	2713
Qy	1021	AAGCATC	1080
Db	2714	AAGCATC	2773
Qy	1081	GCCAAAC	1140
Db	2774	GCCAAAC	2833
Qy	1141	ACCCTCT	1200
Db	2834	ACCCTCT	2893
Qy	1201	ATTGGTG	1260
Db	2894	ATTGGTG	2953
Qy	1261	GTCGGGA	1320
Db	2954	GTCGGGA	3013
Qy	1321	GGAGAC	1380
Db	3014	GGAGAC	3073
Qy	1381	GACCTGA	1440

Db	3074	GACCTGA	3133
Qy	1441	CGTGCCAT	1500
Db	3134	CGTGCCAT	3193
RESULT 11			
US-09-962-436-562			
; Sequence 562, Application US/09962436			
; Patent No. US20020081301A1			
; GENERAL INFORMATION:			
; APPLICANT: Soppet, Daniel			
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat			
; TITLE OF INVENTION: Sets			
; FILE REFERENCE: 689290-75			
; CURRENT APPLICATION NUMBER: US/09/962,436			
; PRIOR APPLICATION NUMBER: US/60/235,082			
; PRIOR FILING DATE: 2000-09-25			
; PRIOR APPLICATION NUMBER: US/60/234,924			
; PRIOR FILING DATE: 2000-09-25			
; NUMBER OF SEQ ID NOS: 568			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 562			
; LENGTH: 7878			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-962-436-562			
Query Match 62.1%; Score 1220; DB 3; Length 7878;			
Best Local Similarity 88.8%; Pred. No. 0;			
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;			
Qy	1	ATGACGTG	60
Db	1697	ATGACATG	1756
Qy	61	AAATACGA	120
Db	1757	AAATACGA	1816
Qy	121	AGCTACCA	180
Db	1817	ATCAACAG	1873
Qy	181	ACTGCGGG	240
Db	1874	ACTGACGG	1933
Qy	241	ATGCTGGG	300
Db	1934	ATGCTGGG	1993
Qy	301	GGAATGCC	360
Db	1994	GGAATGCC	2053
Qy	361	AAAGTGA	420
Db	2054	AAAGTGA	2113
Qy	421	CACATCC	480
Db	2114	CACATCC	2173
Qy	481	GATGATA	540
Db	2174	GATGATA	2233
Qy	541	AACCCGA	600
Db	2234	AACCCGA	2293

QY 601 CTTCTGAGAGATGATCTTGGGCACTGGTGACGTCTGGCCAGTGAGGCACTCC 660
Db 2294 CTGCTGAGGAGGAGCGCATTTGGGGCACTGGTGACGTCTGGCCAGTGAGGCACTCC 2353
QY 661 CTGAGGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAGACCAAGAG 720
Db 2354 CTGCAAGATTCCACAGCCCAATGGTGAGACAGTCCAGGGGCTCCAGACCAAGAG 2413
QY 721 CATGTGTAGCCACGTACACCAAGACCATGGGGCATCAGACAAGAAAGATCTATGT 780
Db 2414 CATGTGTAGCCCAAGTACACCAAGACCATGTGGCATCAGACAAGAGGTCTATGC 2473
QY 781 GGGCAGTGTTCCTCCGTTAGGCTCCTCATCCGATATTGATGACGGGATCTCTCGGG 840
Db 2474 GGGCAGTGTGCTCCTGTTAGGCTCCTTCTCCGAACTGATGACGGGATCTCTCGGG 2533
QY 841 CTCACCTTGCGCTGTGGAGCTGTATCTGTGAGAGCGCAACAGCGCTTGATGCCGATA 900
Db 2534 CTCACCTTGCGCTGTGGAGCTGTATTGTTGAGAGAGAACAGGTGTTGATGCCAATA 2593
QY 901 ACAAGATCGCCTTTAAGTTACAGACAGAGCGCTCAAGAAAGCTCCAGGTGGCCCG 960
Db 2594 ACCAGCATTTGCTTTAAGTTACAGACAGCGCTCATGAAAGACATCCAGGTGGCCCTG 2653
QY 961 TGGGCACTTTTTCACACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1020
Db 2654 TGGGCACTGTCTGCGAACCATTCTTCGATACCTGGGCCATGAACGATGACACCGTCTC 2713
QY 1021 AAGCATTTAGGGCTCTATGAGAAACTAACAAAGACAGGGGACCTGCCACCCCA 1080
Db 2714 AAGCATTTAGGGCTCTACGAGAAACTAACAAAGACAGGGGACCTGCCACCCCA 2773
QY 1081 GCCAAACCCGAGCAAGGCTCTGGCATCCAGGCTGTCCGGCTTACGTTGGCGGAAAG 1140
Db 2774 GCCAAACCCGAGCAAGGCTCTGGCAACCAGGCTGTCCGGCTTACGTTGGCGGAAAG 2833
QY 1141 ACCCTCTGCAAGGGGAGCAGGAGGCCCCCTCCAGGCCCAAGCCCGGTTCCCGCGGCC 1200
Db 2834 ACCCTCTGCAAGGGGTATAGCAGGCCCCCTCCAGGCCCAAGCCAGTTCCAGCGGCC 2893
QY 1201 ATTGCTAGCTTCCCGCCACGGGACCTCTGTTCTTCCACACCTGTCTGTGGGCT 1260
Db 2894 ATTGCTAGCTTCCCGCCATGGGCATCTGTTTTCACGCGCTGTCTGTGGGCT 2953
QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGGCTAG 1320
Db 2954 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCTGGCTAG 3013
QY 1321 GAGAGACCTCAGGGTCTCTGAGATTCTCTGCAAGTGAAGTCCATGCCCCGCTCCCAACG 1380
Db 3014 GAGAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAAGTCAATGCCCCGCTCCCAACG 3073
QY 1381 GACCTGAGCTAGAGGGCTTGGTTCGCGCATTTATGATTTCAAGACAGAGCTGTGGGTG 1440
Db 3074 GACCTGAGTATAGGGGCTTGGTTCGCGCATTTATGATTTGAACGAGCTGTGGGTG 3133
QY 1441 CGTGCCATATCCAGAGGAGACAGCTGGCCCCCTGTGGCAGGCTGAACACCTGGCGAG 1500
Db 3134 CGTGCCATATCCAGAGGAGACAGCTGGCCACTGCTGGCAGGCTGAACACTGGCGAGAG 3193

RESULT 12
US-10-843-641A-3021
; Sequence 3021, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
; CURRENT FILING DATE: 2004-05-12

; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3021
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-843-641A-3021

Query Match 62.1%; Score 1220; DB 9; Length 7878;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 1 ATGAGCGTGTAGAGTCCGGGCAAGTTGTGGGCAACAGCGGAGCATCATATATG 60
Db 1697 ATGAGCATGTAGAGATGCAATAGATTGTGAGGCAAGAGCGGAGCATATCTATG 1756
QY 61 AATACGAAAGGAGCAGCGAGCTGGGCTGCCAGAGCAAGGGGCTTAACCTTTTGA 120
Db 1757 AAGTATGACAGGAGCAGCGAGCTGGGCTGCCAGAGCAAGGGGCTGAACCTTTGA 1816
QY 121 AGTACAACAACAAGCTGATCATTTGGGATTTGTACATGAGACGAGCTGCTCTG 180
Db 1817 ATCAACAGCAGCA---TTGATCGTTTGGCATTTTGACATGAGACGAGCTGCTCTG 1873
QY 181 ACTCGCGGAGGCGCAACAATTCCGCGGAGATCAGCCGAAGCAAGTGGTGAT 240
Db 1874 ACTGACGCGGAGGCGCAAGAAATTCGCGGAGATGACACGAACGAGCAAGTGGATGAA 1933
QY 241 ATGCTGGAGACTGGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCTACAAG 300
Db 1934 ATGCTGGAGAGATGGAGACATATTAAGCAGACATCAAACTCATAGATCGAGTACAAAG 1993
QY 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACATGAGGAATG 360
Db 1994 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGAACATGAGGAATG 2053
QY 361 AAGTGAAGAAACCCGGAAGATACCAAGATGAAGAGAGGGAAGGATCATCTGAG 420
Db 2054 AAGTGAAGAAACCCGGAAGATACCAAGATGAAGAGAGGGAAGGATCATCTGAA 2113
QY 421 CACATCCAGCGCATGACCCGGGACGTAAGCCGGACATTAAAGAAAGATATTCTTACG 480
Db 2114 CACATCCAGCACATGACCTGAGCTGAGAGCACTCTCCGGAACATGTCTTCTTACG 2173
QY 481 GATCGATACGAAACAGCAGCGGGAATCTTCCACATCTCTCTGACATATGAGAGAT 540
Db 2174 GATCGATATGAGCCAGCAGAGGGAATTTCTACATCTCTCTGACATATGAGAGAT 2233
QY 541 AACCCGAGGTGGGCTACTGCAAGGAGCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 2234 AACCCGAGGTGGGCTACTGCAAGGAGCTGAGCCACATCAACCGCTTGTCTCTCTAT 2293

QY 601 CTTCTGAGGAGATGCAATTCTGGGCACTGGTGCAAGCTGCTGGCCAGTGAGAGCACTCC 660
|||
Db 2294 CTGCTGAGAGAGACGCAATCTGGGCACTGGTGCAAGCTGCTGGCCAGTGAGAGCACTCC 2353
QY 661 CTGACGGATTTCACAGCCCAATGGCGGGAACCGTCCAGGGGCTCCAAGACCAAGAG 720
|||
Db 2354 CTGCCAGGATTCCACAGCCCAATGGTGGAAGTCCAGGGGCTCCAAGACCAAGAGAG 2413
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGCAATGAGCAAGAAGAATCTATGT 780
|||
Db 2414 CATGTGTAGCCCAAGTCAACAACCAAGACCATGTGGCATGAGACAGGAAGGTCTATGC 2473
QY 781 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
|||
Db 2474 GGGCAGTGTGCTCTGTTAGGCTGCTCTCCGGAACCTGATTGACGGGATCTCTCGGG 2533
QY 841 CTCACCTGCGGCTGTGGGACGTGTATCTGTGAAAGGCGAACAGCGTTGATGCCGATA 900
Db 2534 CTCACCTGCGGCTGTGGGACGTGTATTTGTGGAAGAGAACAGGTGTGATGCCAATA 2593
QY 901 ACAAGAAATCGCCTTTAAGTTTCAAGAGAAGCGCTCACGAGAAGCTCCAGGTGCGCCG 960
|||
Db 2594 ACCAGCAATGCTCTTAAGTTTCAAGAGAAGCGCTCATGAGAAGATCCAGGTGCGCCG 2653
QY 961 TGGGACGTTTTTGCAACCGGTTGTTGATACCTGGGCCAGGAGTGAAGACACTGTCTC 1020
|||
Db 2654 TGGGACGTTCTGCGGAACCAATCTCGATACCTGGGCCATGAACGATGACACCGTCTC 2713
QY 1021 AAGCATTTAAGGCTCTATGAGAACTAAACAAGAAAGAGGGGACCTGCCACCCCA 1080
|||
Db 2714 AAGCATTTAAGGCTCTACGAGAACTAAACAAGAAAGAGGGGACCTGCCACCCCA 2773
QY 1081 GCCAAACCGGAGCAAGGGTGTGCGGATCCAGGCTGTGCGGCTTCAAGTGCGGGAAG 1140
|||
Db 2774 GCCAAACCGGAGCAAGGGTCTTGCGACCCAGGCTGTGCGGCTTCAAGTGCGGGAAG 2833
QY 1141 ACCCTCTGCAAGGGGACAGGAGCGCCCTCCAGGCCCCACAGCCCGGTTCCCGGGCC 1200
|||
Db 2834 ACCCTCTGCAAGGGGTATAGGAGCGCCCTCCAGGCCCCACAGGTTCCAGGGGCC 2893
QY 1201 ATTTGGTCACTTCCCGGCAAGGCACTGCTTCTTCCACACCTGCTGCTGAGGGCT 1260
|||
Db 2894 ATTTGGTCACTTCCCGGCAAGGCACTGCTTCTTCCACAGCCCTGCTGAGGGCT 2953
QY 1261 GTCCGGAGAACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
|||
Db 2954 GTCCGGAGAACACGTAACCTGTGGGCACTCAGGGTGTGCCAGCCCTGCTCAG 3013
QY 1321 GGAGACTTCAAGGGTCTCTGGAATCTCTGCACTGGAATTCATGCCCCGCTCCCAAG 1380
|||
Db 3014 GGAGACTTCAAGGGTCTCTGGAATCTCTGGAATTCATGCCCCGCTCCCAAG 3073
QY 1381 GACCTGAGTAGAGGGCCCTTGGTTCCGCCATTATGATTTCAGACAGAGCTGCTGGGTC 1440
|||
Db 3074 GACCTGAGTAGAGGGCCCTTGGTTCCGCCATTATGATTTCAGACAGAGCTGCTGGGTC 3133
QY 1441 CGTGCCATATCCAGAGAGACCAAGCTGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1500
|||
Db 3134 CGTGCCATATCCAGAGAGACCAAGCTGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 3193
RESULT 13
US-10-956-157-1259
; Sequence 1259, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04

; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 1259
; LENGTH: 7878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1259
Query Match 62.1%; Score 1220; DB 9; Length 7878;
Best Local Similarity 88.8%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY 1 ATGACGTGTAGAGTCCGGGCAAGTTGTGGGCAACAAGCGAGAGCATCATATG 60
|||
Db 1697 ATGACATGTGTAGAGATGAGATAGTTTGCAGGACAGAGCGGAAGACATACTATG 1756
QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCTTAACCTTTTGA 120
|||
Db 1757 AAGTATGACAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCTTGAAGCCGTTGA 1816
QY 121 AGCTACAAACAACGTCATCATTTGGGGATTGTATCATGAGACGAGCTGCTCTCTG 180
|||
Db 1817 ATCAACAGACGA--TTGATCGTTTGGCATTTTGCATGAGACGAGCTGCTCTCTG 1873
QY 181 ACTGCGGGGAGCGGAAGCAATTCGCGGAGATCAGCCGAAGAGCAAGTGGTGAT 240
|||
Db 1874 ACTGACGGGAGCGGAAGAAATTCGCGGAGATGACAGAACGAGCAAGTGGATGAA 1933
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAGAAAAGCTCATAGATCGAGCTACAG 300
|||
Db 1934 ATGCTGGAGAGATGGAGACATATAAGCACAGTAGCAAACTCATAGATCGAGTACAA 1993
QY 301 GGAATCCCATGAACATCCGGGGCCCGATGTGTCACTCTCTCTGAACATTGAGGAATG 360
|||
Db 1994 GGAATCCCATGAACATCCGGGGCCCGGTGTGTCACTCTCTCTGAACATTGAGGAATC 2053
QY 361 AAGTTGAAAAAACCCTGGAGATACAGATCATGAAGAGAGGCGAAGAGTCACTGAG 420
|||
Db 2054 AAGTTGAAAAAACCCTGGAGATACAGATCATGAAGAGAGGCGAAGAGTCACTGAA 2113
QY 421 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTAAAGAAAGCATATATTCTTCAG 480
|||
Db 2114 CACATCCAGCATGACCTGAGCGTAGAGACGACTCTCCGMAACATGTCTTTAGG 2173
QY 481 GATCGATACGGAACCAAGCAGCGGAAGTACTCCACATCTCTGCGCATATGAGAGTAT 540
|||
Db 2174 GATCGATATGAGAGCCAGAGAGGAGACTATTTACATCTCTGCGCATATTCGAGTAT 2233
QY 541 AACCCGAGGTGGGCTACTGAGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
|||
Db 2234 AACCCGAGGTGGGCTACTGAGGGACCTGAGCCACATCACCGCTTGTCTCTCTAT 2293
QY 601 CTTCTGAGGAGGATGCAATCTGGGCACTGTGTGAGCTGTGGCCAGTGAAGGCACTCC 660
|||
Db 2294 CTGCTGAGGAGGACGCAATCTGGGCACTGTGTGAGCTGTGGCCAGTGAAGGCACTCC 2353
QY 661 CTGACGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 720
|||
Db 2354 CTGCCAGGATTCCACAGCCCAATGGTGGAAGTCCAGGGGCTCCAAGACCAAGAGAG 2413
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGCAATGAGCAAGAAGAATCTATGT 780
|||
Db 2414 CATGTGTAGCCCAAGTCAACAACCAAGACCATGTGGCATCAGAGCAAGGAAGTCTATGC 2473
QY 781 GGGCAGTGTCCCGCTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
|||
Db 2474 GGGCAGTGTGCTCTGTTAGGCTGCTCTCCGGAACCTGATTGACGGGATCTCTCGGG 2533
QY 841 CTCACCTGCGGCTGTGGGACGTGTATCTGTGAAAGGCGAACAGCGCTTGAAGCCGATA 900
|||
Db 2534 CTCACCTGCGGCTGTGGGACGTGTATTTGTGGAAGAGAAACAAGGTGTGATGCCAATA 2593
QY 901 ACAAGAAATCGCCTTTAAGTTTCAAGAGAAGCGCTCACGAAAGCTCCAGGTGTGCGCG 960

Db 2594 ACCAGCATTTGCTTTAAGGTTCAAGCAGAACGCCCTCATGAAGACATCCAGGTGTGGCTTG 2653
QY 961 TGGGCAAGTTTTCGAACCGGTTCTGTTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC 1020
Db 2654 TGGGCAAGTTTTCGAACCGGTTCTGTTGATACCTGGGCCAGGAGTAGAGACACTGTGCTC 2713
QY 1021 AAGCATCTTAGGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGAACTTGCCACCCTCA 1080
Db 2714 AAGCATCTTAGGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGAACTTGCCACCCTCA 2773
QY 1081 GCCAAACCGGACGAAAGGTCGTGGGCAATCCAGGCTGTGCCGGCTTCAAGTGGCGGGAAG 1140
Db 2774 GCCAAACCGGACGAAAGGTCGTGGGCAATCCAGGCTGTGCCGGCTTCAAGTGGGGAAG 2833
QY 1141 ACCCTCTGCAAGGGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 2834 ACCCTCTGCAAGGGGATATAGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2893
QY 1201 ATTTGTCAGCTTCCCCCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 2894 ATTTGTCAGCTTCCCCCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2953
QY 1261 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGGTGTGCCAGGAGGAGGAGGAGGAGGAG 1320
Db 2954 GTCCGGGAAGACACTACCTGTGGGCACTCAGGGGTGTGCCAGGAGGAGGAGGAGGAGGAG 3013
QY 1321 GGAGACCTCAGGGGTCCTGAGATTCCTGAGTGAATGCCATGCCCGGCTCCCAAG 1380
Db 3014 GGAGACCTCAGGGGTCCTGAGATTCCTGAGTGAATGCCATGCCCGGCTCCCAAG 3073
QY 1381 GACCTGAGCTAGAGGGGCTTGTGTCGCCCATTTATGATTTTCAAGACAGAGCTGTGGGTC 1440
Db 3074 GACCTGAGCTAGAGGGGCTTGTGTCGCCCATTTATGATTTTGAACGGAAGCTGTGGGTC 3133
QY 1441 CGTGCCATATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1500
Db 3134 CGTGCCATATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3193

RESULT 14

US-10-450-763-24016
; Sequence 24016, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 24016
; LENGTH: 8180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (4369)..(6126)
; OTHER INFORMATION: 97% homologous to Homo sapiens oncogene, accession number
; OTHER INFORMATION: X63546, Smith-Waterman Score=3063.
US-10-450-763-24016

Query Match 50.9%; Score 999; DB 9; Length 8180;
Best Local Similarity 77.8%; Pred. No. 3.3e-305;
Matches 1332; Conservative 0; Mismatches 165; Indels 214; Gaps 2;

QY 1 ATGACGTGTAGAGTTCGCGGCACTTGTGGCAACAAGCGAGAGACATCATATATG 60
Db 1697 ATGACATGTTAGAGATGCAAGATGTTTGACAGGCACAGAGCGGAAGACATACATATATG 1756
QY 61 AAATACGAAAGGACACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCCA 120
Db 1757 AAGTATGACAAAGGACACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCCA 1816
QY 121 AGCTACAAACAACGTCGATCATTTGGGGAATTGTACATGACAGGAGCTGCTCTCTG 180
Db 1817 A--TCAACAGCAGCATTGATCGTTTGGCAATTTTGCAATGAGACGAGCTGCTCTCTG 1873
QY 181 ACTCGCGGAGCGCAAGCAAAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240
Db 1874 ACTCACGGAGCGCAAGCAAAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 1933
QY 241 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGTCGAGCGTACAG 300
Db 1934 ATGCTGGAGAAATGGGAGACATATACACAGTAGCAAACTCATAGTCGAGTACAG 1993
QY 301 GGAATGCCATGAACATCCGGGCGCGATGTGTGTCAGTCTCTGAACATTTGAGAAATG 360
Db 1994 GGAATGCCATGAACATCCGGGCGCGATGTGTGTCAGTCTCTGAACATTTGAGAAATC 2053
QY 361 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCAATGAG 420
Db 2054 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCAATGAG 2113
QY 421 CACATCCAGCGCATGCACCGGAGCTAAGCGGACATTAAGAAACATATATTTTCAGG 480
Db 2114 CACATCCAGCATGCACCGGAGCTAAGCGGAGCTAAGCGGACATTAAGAAACATATTTTCAGG 2173
QY 481 GATGATACGAAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 540
Db 2174 GATGATACGAAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 2233
QY 541 AACCAGGAGTGGGCTACTGACGGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 600
Db 2234 AACCAGGAGTGGGCTACTGACGGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 2293
QY 601 CTTCTGAGAGGATGCATTTCTGGGCACTGTGACGTGCTGGCCAGTGAGGCACTCC 660
Db 2294 CTTCTGAGAGGAGCGCATTTCTGGGCACTGTGACGTGCTGGCCAGTGAGGCACTCC 2353
QY 661 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAGAG 720
Db 2354 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAGAG 2413
QY 721 CATGTGTAGCCAGTCAAGCCCAAGCCATGAGGGGCAATCAGCAAGAAAGATCTATGT 780
Db 2414 CATGTGTAGCCAGTCAAGCCCAAGCCATGAGGGGCAATCAGCAAGAAAGATCTATGC 2473
QY 781 GGGCAGTGTCCCGGTTAGGCTGCTCATCCGGATATTGATTTGACGGGATCTCTCGGG 840
Db 2474 GGGCAGTGTCCCGGTTAGGCTGCTCATCCGGGATATTGATTTGACGGGATCTCTCGGG 2533
QY 841 CTCAACCTGCGGCTGTGGAGCTGTATCTGTGTAAGGCGAAACAGGCGTTGATGCCGATA 900
Db 2534 CTCAACCTGCGGCTGTGGAGCTGTATTTGTTGTAAGGAGAAACAGGCGTTGATGCCAATA 2593
QY 901 ACAAGAAATCGCCTTTAAGTTCAAGAGCGGCTCAAGAAAGCTCCAGGTGTGGCCCG 960
Db 2594 ACCAGCATGCTCTTAAAGTTCAAGAGCGGCTCAAGAAAGCTCCAGGTGTGGCCCG 2653
QY 961 TGGGCAAGTTTTCGAACCGGTTCTGTATACCTGGGCAAGGATGAGACACTGTGCTC 1020
Db 2654 TGGGCAAGTTTTCGAACCGGTTCTGTATACCTGGGCAAGGATGAGACACTGTGCTC 2713
QY 1021 AAGCATCTTAGGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGAGCTGCAACCC-- 1078
Db 2714 AAGCATCTTAGGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGAGCTGCAACCCCA 2773
QY 1079 ----- 1078

Db	2774	GGCCCAACAGCCCTGGGACGAAGGTGTGTGGCAGGAAGCCCCCAGCCAGCTGTAACCTTG	2833
QY	1079	-----	1078
Db	2834	GGGGCAGTCCCAAGAGCCACCACCATGCCCCAACGGCTTCCCATGCCAGGCAGCACAC	2893
QY	1079	-----	1078
Db	2894	ACCCCTCCCTGGGATCAGCAGACTACAGGCGTGTCTCATGTCTCAGACCAAGGGCC	2953
QY	1079	-----CAGCCAAACCCGACGAAGGTGTCGGCATC	1109
Db	2954	ACACAGAGACCCCAAGACTCCAGAGATGCAGCCAAACGGAGCAAGGTCTTGCAACC	3013
QY	1110	CAGGCTGTGCGGGCTTACGTCGCGGGAAGACCCTCTGCAAGGGGACAGGCAGGCCCC	1169
Db	3014	CAGGCTGTGCGGCTTACGTCGTCGGAAGACCCTCTGCAAGGGGTATAGGCAGGCCCC	3073
QY	1170	TCCAGGCCCCACAGCCCGGTTCCCGCGGCCCATTTGTCAGCTTCCCCGCACGGGCACC	1229
Db	3074	TCCAGGCCCCACAGCCCGAGTTCCAGCGGCCCATTTGCTCAGCTTCCCCGCATGGGCATC	3133
QY	1230	TGCTTCTTCCACACCCCTGTCTGTGTGGGCTGTCCGGGAAGACCTACCCCTGTGGCAC	1289
Db	3134	TGCTTTTCCACGCCCTGTCTGTGTGGGCTGTCCGGGAAGACCTACCCCTGTGGCAC	3193
QY	1290	TCAGGGTGTCCCAAGCCCGCCCTGGCTCAGGAGGACCTCAGGGTTCCTGAGATTCT	1349
Db	3194	TCAGGGTGTCCCAAGCCCTGGCTCAGGAGGACCTCAGGGTTCCTGAGATTCT	3253
QY	1350	GCAGTGGACTCCATGCCCCGCTCCCAACGGAACCTGACGTAGAGGGCCCTTGTTCCG	1409
Db	3254	GGAAGTGAAGTCAATGCCCCGCTCCCAACGGAACCTGATATAGGGGGCCCTTGTTCC	3313
QY	1410	CCATTATGATTTCAAGACAGAGCTGTGGTCCGTGCCATATCCACGAGGACCAAGTGGC	1469
Db	3314	CCATTATGATTTTGAACGAGAGCTGTGGTCCGTGCCATATCCACGAGGACCAAGTGGC	3373
QY	1470	CCCCTGCTGGCAGGCTGAACACCTTGCGGAG	1500
Db	3374	CACCTGTGGCAGGCTGAACACTGCGGAGAG	3404

RESULT 15

```

; Sequence 272, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 272
; LENGTH: 8201
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-272

```

Query Match	50.9%	Score 999;	DB 7;	Length 8201;
Best Local Similarity	77.8%	Pred. No. 3.3e-305;		
Matches 1332; Conservative	0;	Mismatches 165;	Indels 214;	Gaps 2;

QY	1	ATGACGCTGTAGAGGTGCGGGCAGTTGGTGGGCAAGAGCGGAGGACATCATTAATG	60
Db	1697	ATGACATGTGTAGAAATGCAGATAGTTTGCAGGCACAGAGCGGAAGGACATACTTAATG	1756
QY	61	AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA	120
Db	1757	AAGTATGACAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTGAGCCCGTTTGA	1816
QY	121	AGCTACAACAACGTCGATCATTTGGGATTTGTACATGAGCGAGCTGCTCTCTG	180
Db	1817	A---TCAACAGCAGACGATTGATCGTTTGGCATTTTGATGAGACGAGCTGCCCTGTG	1873
QY	181	ACTCGCGGGAGGCGGAACCAATTCCGCGGGAGATCAAGCCGAAAGCAAGTGGTGGAT	240
Db	1874	ACTGACCGGAGGCGAAGAAAATTCCGCGGGAGATGACACGAACGAACTGATGGAA	1933
QY	241	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAAGCTCATAGATGAGCGTACAAG	300
Db	1934	ATGCTGGGAGAAATGGAGACATTAAGCACAGTACAAACTCATAGATGAGTGTACAAG	1993
QY	301	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTCCTGAACATTGAGAAATG	360
Db	1994	GGAATTCCTCATGAACATCCGGGGCCCGGTGTGTCACTCTCCTGAACATTGAGAAATC	2053
QY	361	AAGTTGAAAAACCCCGGAAGATACCAGATCATGAAGAGGAAGGGCAAGAGTCACTGAG	420
Db	2054	AAGTTGAAAAACCCCGGAAGATACCAGATCATGAAGAGAGGGCAAGAGTCACTGAA	2113
QY	421	CACATCCAGCGCATGCAACCGGACGTAAAGCGGGACATTAAGGAAGCATATATCTTCAGG	480
Db	2114	CACATCCACACATGCACTGAGCGTGAAGACGACTCTCCGAAACATGTCTTCTTAAAG	2173
QY	481	GATCGATACGGAACCAAGCAGCGGAACTACTCCATCTCCTGCGATATGAGAGTAT	540
Db	2174	GATCGATATGAGCCAAAGCAGAGGAACTATTCTACATCTCCTGCGCTATTGCGAGTAT	2233
QY	541	AACCCGAGGTGGGCTACTGCAGGAGCCTGAGCCACATCGCCGCTTGTCTCTCTAT	600
Db	2234	AACCCGAGGTGGGCTACTGCAGGAGCCTGAGCCACATCACCGCCTTGTCTCTCTAT	2293
QY	601	CTTCCTGAGGAGATGCATTCTGGGCACTGTGCAGCTGCTGCCAGTGAGGCACTCC	660
Db	2294	CTGCCTGAGGAGACGCATTCTGGGCACTGTGTGCACTGCTGCCAGTGAGGCACTCC	2353
QY	661	CTGCAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG	720
Db	2354	CTGCAGGATTTCAAGCCCAATGTGGGACGTCCAGGGGCTCCAAGACCAAGAG	2413
QY	721	CATGTGTAGCCAGTCACAACCCCAAGACCAATGGGGCATCAGGACCAAGAAAGATCTATGT	780
Db	2414	CATGTGTATCCCAAGTCACAACCCCAAGACCAATGGGCATCAGGACCAAGAAAGTCTATGC	2473
QY	781	GGGAGTGTCCCGCTTAGGCTGCTCATCCGGATATTGATTGACGGGATCTCTCGGG	840
Db	2474	GGGAGTGTCCCTCTGCTTAGGCTGCTCTCTCCGAACTGATTGACGGGATCTCTCGGG	2533
QY	841	CTCACCTTGCGCTGTGGGACGTGTATCTGTAGAAAGCGGAACAGGCGTTGATGCCGATA	900
Db	2534	CTCACCTTGCGCTGTGGGACGTGTATTTGTGGAAGGAAGAACAGGTGTTGATGCCAATA	2593
QY	901	ACAAGAAATCGCCTTTAAGTTTCAAGCAGAAAGCGCTCACGAAGAGTCCAGGTGTGCCCCG	960
Db	2594	ACCAGCATTTGCTTTAAAGTTTCAAGCAGAAAGCGCTCATGAAGACATCCAGGTGTGCCCCG	2653
QY	961	TGGGACGTTTTTGAACCGGTTGTTGATACCTGGGCCAGGGATGAGGACATGTGCTC	1020
Db	2654	TGGGACAGTCTGCGGAACCAATTTCTGATACCTGGGCCATGAACGATGACACGTTGCTC	2713
QY	1021	AAGCATCTTAAGGCTCTATGAAGAACTAACAAAGAACAGGGGCACTGCGCACCCC--	1078
Db	2714	AAGCATCTTAAGGCTCTTAAGAAAGAACTAACAAAGAACAGGGGCACTGCGCACCCCA	2773

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 14:20:23 ; Search time 1321 Seconds
(without alignments)
5947.837 Million cell updates/sec

Title: US-10-071-838-1
Perfect score: 1964
Sequence: 1 atggagctgtagagtcgc.....aatgtctcctgtgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9267905 seqs, 2000278028 residues

Total number of hits satisfying chosen parameters: 18535810

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New: *
1: /SIDS5/ptodata/1/pubpna/US08_NEW_PUB.seq: *
2: /SIDS5/ptodata/1/pubpna/US06_NEW_PUB.seq: *
3: /SIDS5/ptodata/1/pubpna/US07_NEW_PUB.seq: *
4: /SIDS5/ptodata/1/pubpna/PCT_NEW_PUB.seq: *
5: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq: *
6: /SIDS5/ptodata/1/pubpna/US09_NEW_PUB.seq: *
7: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq: *
8: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq: *
9: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq: *
10: /SIDS5/ptodata/1/pubpna/US10_NEW_PUB.seq: *
11: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq: *
12: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq: *
13: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq: *
14: /SIDS5/ptodata/1/pubpna/US11_NEW_PUB.seq: *
15: /SIDS5/ptodata/1/pubpna/US60_NEW_PUB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	358.6	18.3	378	6	US-09-925-065A-506682, Sequence 506682,
C 2	358.6	18.3	378	6	US-09-925-065A-506683, Sequence 506683,
C 3	301.6	15.4	186442	14	US-11-121-086-104, Sequence 104, App
C 4	300.6	15.3	477	6	US-09-925-065A-198638, Sequence 198638,
C 5	300.2	15.3	477	6	US-09-925-065A-198636, Sequence 198636,
C 6	300.2	15.3	477	6	US-09-925-065A-198637, Sequence 198637,
C 7	300.2	15.3	477	10	US-10-301-480-287418, Sequence 287418,
C 8	300.2	15.3	477	10	US-10-301-480-900827, Sequence 900827,
C 9	297.2	15.1	468	6	US-09-925-065A-173389, Sequence 173389,
C 10	297.2	15.1	477	10	US-10-301-480-264964, Sequence 264964,
C 11	297.2	15.1	477	10	US-10-301-480-878373, Sequence 878373,
C 12	294.2	15.0	476	10	US-10-301-480-285677, Sequence 285677,
C 13	294.2	15.0	476	10	US-10-301-480-899086, Sequence 899086,
C 14	294	15.0	479	6	US-09-925-065A-201701, Sequence 201701,
C 15	293.8	15.0	469	6	US-09-925-065A-198635, Sequence 198635,
C 16	293.8	15.0	469	10	US-10-301-480-287417, Sequence 287417,
C 17	293.8	15.0	469	10	US-10-301-480-900826, Sequence 900826,
C 18	293.8	15.0	479	10	US-10-301-480-294544, Sequence 294544,

C 19	293.8	15.0	479	10	US-10-301-480-294545, Sequence 294545,
C 20	293.8	15.0	479	10	US-10-301-480-907953, Sequence 907953,
C 21	293.8	15.0	479	10	US-10-301-480-907954, Sequence 907954,
C 22	293.6	14.9	459	10	US-10-301-480-264965, Sequence 264965,
C 23	293.6	14.9	459	10	US-10-301-480-878374, Sequence 878374,
C 24	293	14.9	465	6	US-09-925-065A-173390, Sequence 173390,
C 25	292.6	14.9	465	6	US-09-925-065A-173388, Sequence 173388,
C 26	292.6	14.9	479	10	US-10-301-480-294543, Sequence 294543,
C 27	292.6	14.9	479	10	US-10-301-480-907952, Sequence 907952,
C 28	292.2	14.9	477	10	US-10-301-480-294547, Sequence 294547,
C 29	292.2	14.9	477	10	US-10-301-480-907956, Sequence 907956,
C 30	289.2	14.7	472	6	US-09-925-065A-198614, Sequence 198614,
C 31	287	14.6	479	10	US-10-301-480-290201, Sequence 290201,
C 32	287	14.6	479	10	US-10-301-480-290202, Sequence 290202,
C 33	287	14.6	479	10	US-10-301-480-903610, Sequence 903610,
C 34	287	14.6	479	10	US-10-301-480-903611, Sequence 903611,
C 35	285.8	14.6	475	6	US-09-925-065A-585252, Sequence 585252,
C 36	285.4	14.5	479	10	US-10-301-480-290200, Sequence 290200,
C 37	285.4	14.5	479	10	US-10-301-480-903609, Sequence 903609,
C 38	281.8	14.3	467	6	US-09-925-065A-206786, Sequence 206786,
C 39	281.8	14.3	467	6	US-09-925-065A-206787, Sequence 206787,
C 40	281	14.3	473	6	US-09-925-065A-201700, Sequence 201700,
C 41	281	14.3	473	6	US-09-925-065A-201702, Sequence 201702,
C 42	280.6	14.3	467	6	US-09-925-065A-206784, Sequence 206784,
C 43	280.6	14.3	467	6	US-09-925-065A-206785, Sequence 206785,
C 44	280.2	14.3	465	6	US-09-925-065A-118036, Sequence 118036,
C 45	280.2	14.3	465	6	US-09-925-065A-206789, Sequence 206789,

ALIGNMENTS

RESULT 1
US-09-925-065A-506682/c
; Sequence 506682, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108627.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243, 096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252, 147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250, 092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261, 766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289, 846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 506682
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-506682

Query Match 18.3%; Score 358.6; DB 6; Length 378;
Best Local Similarity 99.7%; Pred. No. 3.8e-89;
Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1079 CAGCCAAACCCGAGCAAGGCTCGTCGATCCAGGCTGTGCGGCTTCAAGTGGCGGA 1138
DB 359 CAGCCAAACCCGAGCAAGGCTCGTCGATCCAGGCTGTGCGGCTTCAAGTGGCGGA 300
QY 1139 AGACCTCTGCAAGGGGAGCAGGAGCCCTCCAGGCCACCAAGCCGGTCCCGCGGC 1198
DB 299 AGACCTCTGCAAGGGGAGCAGGAGCCCTCCAGGCCACCAAGCCGGTCCCGCGGC 240

OY 1199 CCATTGGTACGCTTCCCCCGGACGGGCACTCGTCTTCCACACCCCTGTCTGTGGG 1258
|
DB 239 CCATTGGTACGCTTCCCCCGGACGGGCACTCGTCTTCCACACCCCTGTCTGTGGG 180

OY 1259 CTGTCCGGAGACACCTACCCCTGTGGGCACTAGGGTGTGCCAGCCCGGCTGTCTC 1318
|
DB 179 CTGTCCGGAGACACCTACCCCTGTGGGCACTAGGGTGTGCCAGCCCGGCTGTCTC 120

OY 1319 AGGAGACCTCAGGGTCTCTGAGATTCTGCAGTGAATCCATGCCCGCTCCCA 1378
|
DB 119 AGGAGACCTCAGGGTCTCTGAGATTCTGCAGTGAATCCATGCCCGCTCCCA 60

OY 1379 CGGACCTGACGTAGAGGGCCCTTGGTCCGCATTATGATTTCAGACAGAGCTGTGG 1437
|
DB 59 CGGACCTGACGTAGAGGGCCCTTGGTCCGCATTATGATTTCAGACAGAGCTGTGG 1

RESULT 2

US-09-925-065A-506683/c
; Sequence 506683, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 506683
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-506683

Query Match 18.3%; Score 358.6; DB 6; Length 378;
Best Local Similarity 99.7%; Pred. No. 3.8e-89;
Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1079 CAGCCAAACCCGACGAGGGTGTGCGCATCCAGGCTGTGCCGCTTCACTGCGGGA 1138
|
DB 359 CAGCCAAACCCGACGAGGGTGTGCGCATCCAGGCTGTGCCGCTTCACTGCGGGA 300

OY 1139 AGACCTCTGCAAGGGGGAACAGGAGGCCCCCTCCAGGCCACGCGGTTCCCGCGC 1198
|
DB 299 AGACCTCTGCAAGGGGGAACAGGAGGCCCCCTCCAGGCCACGCGGTTCCCGCGC 240

OY 1199 CCATTGGTACGCTTCCCCCGGACGGGCACTCGTCTTCCACACCCCTGTCTGTGGG 1258
|
DB 239 CCATTGGTACGCTTCCCCCGGACGGGCACTCGTCTTCCACACCCCTGTCTGTGGG 180

OY 1259 CTGTCCGGAGACACCTACCCCTGTGGGCACTAGGGTGTGCCAGCCCGGCTGTCTC 1318
|
DB 179 CTGTCCGGAGACACCTACCCCTGTGGGCACTAGGGTGTGCCAGCCCGGCTGTCTC 120

OY 1319 AGGAGACCTCAGGGTCTCTGAGATTCTGCAGTGAATCCATGCCCGCTCCCA 1378
|
DB 119 AGGAGACCTCAGGGTCTCTGAGATTCTGCAGTGAATCCATGCCCGCTCCCA 60

OY 1379 CGGACCTGACGTAGAGGGCCCTTGGTCCGCATTATGATTTCAGACAGAGCTGTGG 1437
|
DB 59 CGGACCTGACGTAGAGGGCCCTTGGTCCGCATTATGATTTCAGACAGAGCTGTGG 1

RESULT 3

US-11-121-086-104
; Sequence 104, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 104
; LENGTH: 186442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-104

Query Match 15.4%; Score 301.6; DB 14; Length 186442;
Best Local Similarity 80.0%; Pred. No. 3.2e-72;
Matches 389; Conservative 0; Mismatches 74; Indels 23; Gaps 2;

OY 1502 GGGTAGATCGGCTTTCGCTGCACCCACGACTGATTCCGACAGGGCACCCCTCAGAG 1561
|
DB 90831 GGACTAGCTGCCCCCTACAGTGCACACAGCACTGCTTCCACCAAGGGCACCCGTTACAG 90890

OY 1562 CTAGGAGCAACAGCCGTGTGCTTCCCACTCAGGGCTTGCCTGCGGCTTCCACTTGG 1621
|
DB 90891 CTAGGAGCTGACAGCAGTGTGCTTCCCACTTGAAGCCTTGGCTTGCACCTTAATTGG 90950

OY 1622 AAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGCTGATA 1681
|
DB 90951 AATGTTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTTATGTTGATA 91010

OY 1682 ATTTCCTAGGCTTAAACAACCCAGCAAGCTTCGCGCTGTTTATTTTGGTTAAAC 1741
|
DB 91011 AGCACCCGAGGTCCCAACAACCCAAAGCTTCGATCCTGTTTATTTTGTAAAC 91070

OY 1742 TTATGAATAATGATTAAGA-----AAGAGTGACGTCGAGAGA 1779
|
DB 91071 TTATGAATAATTAATTAAGATATGACTTACATCCCTGCATGTAAGTTCAGCTCAAAAGT 91130

OY 1780 GATTCAAGATGGAACACACCAAGCCCCAGATCACAAGCCAAACCATGCCCAAGCCCTCC 1839
|
DB 91131 CATTCAAGATGGAACACACCAAGCCCCAGATCACAAGCCAAACCATGCCCAAGCCCTCC 91190

OY 1840 CAGCACCCCAAGCCCCAGACCATGTTCTGAATTCTGACGACACACCGTAGCCCTGCC-TT 1898
|
DB 91191 CAGTGCCCTAGTCTCTGCAACCACTGTCTGAATTCTGACGACATTTGTAGAGCTGCTTT 91250

OY 1899 TGTACTTTAACTCATGGAAGATACTACCTTCAAGCTTTTGAATAAATGTTTCTGT 1958
|
DB 91251 TGTACTTTAACAATCATGGAAGATACTACCTTCAAGCTTTTGAATAAATGTTTCTGT 91310

OY 1959 GAAATG 1964
|
DB 91311 GAAATG 91316

RESULT 4

US-09-925-065A-198638
; Sequence 198638, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135


```
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198638
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-198638

Query Match          15.3%; Score 300.6; DB 6; Length 477;
Best Local Similarity 98.7%; Pred. No. 5.3e-73;
Matches 303; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1658 TGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGG 1717
    |||||
Db 1 TGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGCA 60

QY 1718 TCCTCGTTTATTTTGGTTAACTTATGAAGAGAGTGCAGCTCGAGA 1777
    |||||
Db 61 TCCTCGTTTATTTTGGTTAACTTATGAAGAGAGTGCAGCTCGAGA 120

QY 1778 GAGATTCAAGATGGAACACACCAAGCCAGATCACAAGCCACCATGCCAGCCCT 1837
    |||||
Db 121 GAGATTCAAGATGGAACACACCAAGCCAGATCACAAGCCACCATGCCAGCCCT 180

QY 1838 CCCAGCACCCTCCAGCCCGACGACCATGTTCTGAATTCGACGACACCGTGAGCCTGCT 1897
    |||||
Db 181 CCCAGCACCCTCCAGCCCGACGACCATGTTCTGAATTCGACGACACCGTGAGCCTGCT 240

QY 1898 TTGACTTTAACTCATGGAAGATACTACCTTCACGTTTGAATAATGTTCTGCT 1957
    |||||
Db 241 TTGACTTTAACTCATGGAAGATACTACCTTCACGTTTGAATAATGTTCTGCT 300

QY 1958 TGAATG 1964
    |||||
Db 301 TGAATG 307

RESULT 5
US-09-925-065A-198636
; Sequence 198636, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198636
```

```
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-198636

Query Match          15.3%; Score 300.2; DB 6; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1658 TGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGG 1717
    |||||
Db 1 TGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGCA 60

QY 1718 TCCTCGTTTATTTTGGTTAACTTATGAAGAGAGTGCAGCTCGAGA 1777
    |||||
Db 61 TCCTCGTTTATTTTGGTTAACTTATGAAGAGAGTGCAGCTCGAGA 120

QY 1778 GAGATTCAAGATGGAACACACCAAGCCAGATCACAAGCCACCATGCCAGCCCT 1837
    |||||
Db 121 GAGATTCAAGATGGAACACACCAAGCCAGATCACAAGCCACCATGCCAGCCCT 180

QY 1838 CCCAGCACCCTCCAGCCCGACGACCATGTTCTGAATTCGACGACACCGTGAGCCTGCT 1897
    |||||
Db 181 CCCAGCACCCTCCAGCCCGACGACCATGTTCTGAATTCGACGACACCGTGAGCCTGCT 240

QY 1898 TTGACTTTAACTCATGGAAGATACTACCTTCACGTTTGAATAATGTTCTGCT 1957
    |||||
Db 241 TTGACTTTAACTCATGGAAGATACTACCTTCACGTTTGAATAATGTTCTGCT 300

QY 1958 TGAATG 1964
    |||||
Db 301 TGAATG 307

RESULT 6
US-09-925-065A-198637
; Sequence 198637, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198637
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-198637

Query Match          15.3%; Score 300.2; DB 6; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1658 TGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGG 1717
    |||||
Db 1 TGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGCA 60

QY 1718 TCCTCGTTTATTTTGGTTAACTTATGAAGAGAGTGCAGCTCGAGA 1777
    |||||
```

```
Db      61 TCCTCGTTTATTTTGGTTAAACTTATGAAATGTATTAAAGAAAGAGTCAGCTCGAGA 120
QY      1778 GAGATTCAAGATGGAACACACACAGACCCCGATGCACAAAGCCACCATGCCAGCCCT 1837
Db      121 GAGATTCAAGATGGAACACACACAGACCCCGATGCACAAAGCCACCATGCCAGCCCT 180
QY      1838 CCCAGCACCCCGACGCCACGACCATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCCCT 1897
Db      181 CSCAGCACCCCGACGCCACGACCATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCCCT 240
QY      1898 TTGTACTTTAAACTCATGGAAGAGATACTACCTTCACGTTTGAATAAATGTTTCCGT 1957
Db      241 TTGTACTTCAAACTCATGGAAGAGATAACCACTTCATGTTTGAATAAATGTTTCCGT 300
QY      1958 TGAATG 1964
Db      301 TGAATG 307
```

```
RESULT 7
US-10-301-480-287418/c
; Sequence 287418, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 287418
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-287418
```

```
Query Match      15.3%; Score 300.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1658 TGGGCCAGGGCTCATGCGCTGATATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTGGCG 1717
Db      477 TGGGCCAGGGCTCATGCGCTGATATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTGGCA 418
QY      1718 TCCTCGTTTATTTTGGTTAAACTTATGAAATGTATTAAAGAAAGAGTCAGCTCGAGA 1777
Db      417 TCCTCGTTTATTTTGGTTAAACTTATGAAATGTATTAAAGAAAGAGTCAGCTCGAGA 358
QY      1778 GAGATTCAAGATGGAACACACACAGACCCCGATGCACAAAGCCACCATGCCAGCCCT 1837
Db      357 GAGATTCAAGATGGAACACACACAGACCCCGATGCACAAAGCCACCATGCCAGCCCT 298
QY      1838 CCCAGCACCCCGACGCCACGACCATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCCCT 1897
Db      297 CSCAGCACCCCGACGCCACGACCATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCCCT 238
QY      1898 TTGTACTTTAAACTCATGGAAGAGATACTACCTTCACGTTTGAATAAATGTTTCCGT 1957
Db      237 TTGTACTTCAAACTCATGGAAGAGATAACCACTTCATGTTTGAATAAATGTTTCCGT 178
QY      1958 TGAATG 1964
Db      177 TGAATG 171
```

RESULT 8
US-10-301-480-900827/c

```
; Sequence 900827, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; TITLE OF INVENTION: in the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 900827
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-900827
```

```
Query Match      15.3%; Score 300.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 6.8e-73;
Matches 302; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1658 TGGGCCAGGGCTCATGCGCTGATATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTGGCG 1717
Db      477 TGGGCCAGGGCTCATGCGCTGATATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTGGCA 418
QY      1718 TCCTCGTTTATTTTGGTTAAACTTATGAAATGTATTAAAGAAAGAGTCAGCTCGAGA 1777
Db      417 TCCTCGTTTATTTTGGTTAAACTTATGAAATGTATTAAAGAAAGAGTCAGCTCGAGA 358
QY      1778 GAGATTCAAGATGGAACACACACAGACCCCGATGCACAAAGCCACCATGCCAGCCCT 1837
Db      357 GAGATTCAAGATGGAACACACACAGACCCCGATGCACAAAGCCACCATGCCAGCCCT 298
QY      1838 CCCAGCACCCCGACGCCACGACCATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCCCT 1897
Db      297 CSCAGCACCCCGACGCCACGACCATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCCCT 238
QY      1898 TTGTACTTTAAACTCATGGAAGAGATACTACCTTCACGTTTGAATAAATGTTTCCGT 1957
Db      237 TTGTACTTCAAACTCATGGAAGAGATAACCACTTCATGTTTGAATAAATGTTTCCGT 178
QY      1958 TGAATG 1964
Db      177 TGAATG 171
```

```
RESULT 9
US-09-925-065A-173389
; Sequence 173389, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 173389
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-173389
```

```
Query Match      15.1%; Score 297.2; DB 6; Length 468;
Best Local Similarity 98.4%; Pred. No. 4.6e-72;
Matches 299; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1661 GCCAGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTCGCGTCC 1720
DB      1   GCCAGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTCGCAATCC 60
```

```
QY      1721 TCGTTTATTTTGGTTAAACTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAG 1780
DB      61 TCGTTTATTTTGGTTAAACTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAG 120
```

```
QY      1781 ATTCAAGATGGAACACACACAGACCCAGATCACAAGCCAACCATGCCAGCCCTCC 1840
DB      121 ATTCAAGATGGAACACACAGACCCAGATCACAAGCCAACCATGCCAGCCCTCC 180
```

```
QY      1841 AGCACCCTCAGCCCCACGACCATCGTTCTGAATCTGACGACACCCGTGAGCCTGCTTTG 1900
DB      181 AGCACCCTCAGCCCCACGACCATCGTTCTGAATCTGACGACACCCGTGAGCCTGCTTTG 240
```

```
QY      1901 TACTTTAACTCATGGAAGGATACTTCACTGCTTTGAAATTAATGTTTCCGTGGA 1960
DB      241 TACTTCAAACTCATGGAAGGATACTTCACTGCTTTGAAATTAATGTTTCCGTGGA 300
```

```
QY      1961 AATG 1964
DB      301 AATG 304
```

RESULT 10

```
US-10-301-480-264964
; Sequence 264964, Application US/10301480
; Publication No. US20060057564A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: In the Human Genome
```

```
; CURRENT REFERENCE: 108827.137
```

```
; CURRENT APPLICATION NUMBER: US/10/301,480
```

```
; PRIOR FILING DATE: 2002-11-21
```

```
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
```

```
; PRIOR APPLICATION NUMBER: US 60/311,695
```

```
; PRIOR FILING DATE: 2001-08-10
```

```
; NUMBER OF SEQ ID NOS: 1226818
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 264964
```

```
; LENGTH: 477
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
```

```
US-10-301-480-264964
```

```
Query Match      15.1%; Score 297.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 4.7e-72;
Matches 299; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1661 GCCAGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTCGCGTCC 1720
DB      1   GCCAGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTCGCATCC 60
```

```
QY      1721 TCGTTTATTTTGGTTAAACTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAG 1780
DB      61 TCGTTTATTTTGGTTAAACTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAG 120
```

```
QY      1781 ATTCAAGATGGAACACACACAGACCCAGATCACAAGCCAACCATGCCAGCCCTCC 1840
DB      121 ATTCAAGATGGAACACACACAGACCCAGATCACAAGCCAACCATGCCAGCCCTCC 180
```

```
QY      1841 AGCACCCTCAGCCCCACGACCATCGTTCTGAATCTGACGACACCCGTGAGCCTGCTTTG 1900
DB      181 AGCACCCTCAGCCCCACGACCATCGTTCTGAATCTGACGACACCCGTGAGCCTGCTTTG 240
```

```
QY      1901 TACTTTAACTCATGGAAGGATACTACCTTCAGCTTTGAAATAAATGTTTCCGTGGA 1960
DB      241 TACTTCAAACTCATGGAAGGATACTACCTTCAGCTTTGAAATAAATGTTTCCGTGGA 300
```

```
QY      1961 AATG 1964
DB      301 AATG 304
```

RESULT 11

```
US-10-301-480-878373
; Sequence 878373, Application US/10301480
; Publication No. US20060057564A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: In the Human Genome
```

```
; CURRENT REFERENCE: 108827.137
```

```
; CURRENT APPLICATION NUMBER: US/10/301,480
```

```
; PRIOR FILING DATE: 2002-11-21
```

```
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
```

```
; PRIOR APPLICATION NUMBER: US 60/311,695
```

```
; PRIOR FILING DATE: 2001-08-10
```

```
; NUMBER OF SEQ ID NOS: 1226818
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 878373
```

```
; LENGTH: 477
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
```

```
US-10-301-480-878373
```

```
Query Match      15.1%; Score 297.2; DB 10; Length 477;
Best Local Similarity 98.4%; Pred. No. 4.7e-72;
Matches 299; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```
QY      1661 GCCAGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTCGCGTCC 1720
DB      1   GCCAGGCTCATGGCTGATATTTCCCTAGGCTTAACAACCCAAAGCAAGCTTCGCATCC 60
```

```
QY      1721 TCGTTTATTTTGGTTAAACTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAG 1780
DB      61 TCGTTTATTTTGGTTAAACTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAG 120
```

```
QY      1781 ATTCAAGATGGAACACACACAGACCCAGATCACAAGCCAACCATGCCAGCCCTCC 1840
DB      121 ATTCAAGATGGAACACACACAGACCCAGATCACAAGCCAACCATGCCAGCCCTCC 180
```

```
QY      1841 AGCACCCTCAGCCCCACGACCATCGTTCTGAATCTGACGACACCCGTGAGCCTGCTTTG 1900
DB      181 AGCACCCTCAGCCCCACGACCATCGTTCTGAATCTGACGACACCCGTGAGCCTGCTTTG 240
```

```
QY      1901 TACTTTAACTCATGGAAGGATACTACCTTCAGCTTTGAAATAAATGTTTCCGTGGA 1960
DB      241 TACTTCAAACTCATGGAAGGATACTACCTTCAGCTTTGAAATAAATGTTTCCGTGGA 300
```

```
QY      1961 AATG 1964
DB      301 AATG 304
```

RESULT 12

```
US-10-301-480-285677/c
; Sequence 285677, Application US/10301480
; Publication No. US20060057564A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
```

```

; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285677
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-285677
```

```

Query Match          15.0%; Score 294.2; DB 10; Length 476;
Best Local Similarity 98.3%; Pred. No. 3.2e-71;
Matches 296; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```

QY 1664 AGGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCTCTCG 1723
    |||||||
DB 476 AGGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCTCTCG 417

QY 1724 TTTATTTTGGTTAACTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATT 1783
    |||||||
DB 416 TTTATTTTGGTTAACTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATT 357

QY 1784 CAGAGATGGAACACACAGACCCAGATCAAAAGCCCAACCATGCCAGCCCTCCAGC 1843
    |||||||
DB 356 CAGAGATGGAACACACAGACCCAGATCAAAAGCCCAACCATGCCAGCCCTCCAGC 297

QY 1844 ACCCCAGCCCCACGACCATCGTCTGAATTTGACGACACCGTGAGCTTGCTTTGTAC 1903
    |||||||
DB 296 ACCCCAGCCCCACGACCATCGTCTGAATTTGACGACACCGTGAGCTTGCTTTGTAC 237

QY 1904 TTTAACTCATGGAAGATTAAGTACCTTCACGTTTGAATAATGTTCCTGTGAAAT 1963
    |||||||
DB 236 TTCAACTCATGGAAGATTAAGTACCTTCACGTTTGAATAATGTTCCTGTGAAAT 177

QY 1964 G 1964
    |
DB 176 G 176
```

RESULT 13
US-10-301-480-899086/c

```

; Sequence 899086, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 899086
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-899086
```

```

Query Match          15.0%; Score 294.2; DB 10; Length 476;
Best Local Similarity 98.3%; Pred. No. 3.2e-71;
Matches 296; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

QY 1664 AGGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCTCTCG 1723

```

DB 476 AGGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCTCTCG 417
    |||||||
QY 1724 TTTATTTTGGTTAACTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATT 1783
    |||||||
DB 416 TTTATTTTGGTTAACTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATT 357

QY 1784 CAGAGATGGAACACACAGACCCAGATCACAAGCCCAACCATGCCAGCCCTCCAGC 1843
    |||||||
DB 356 CAGAGATGGAACACACAGACCCAGATCACAAGCCCAACCATGCCAGCCCTCCAGC 297

QY 1844 ACCCCAGCCCCACGACCATCGTTCGAATTTGACGACACCGTGAGCTTGCTTTGTAC 1903
    |||||||
DB 296 ACCCCAGCCCCACGACCATCGTTCGAATTTGACGACACCGTGAGCTTGCTTTGTAC 237

QY 1904 TTTAACTCATGGAAGATTAAGTACCTTCACGTTTGAATAATGTTCCTGTGAAAT 1963
    |||||||
DB 236 TTCAACTCATGGAAGATTAAGTACCTTCACGTTTGAATAATGTTCCTGTGAAAT 177

QY 1964 G 1964
    |
DB 176 G 176
```

RESULT 14
US-09-925-065A-201701/c

```

; Sequence 201701, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201701
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-201701
```

```

Query Match          15.0%; Score 294; DB 6; Length 479;
Best Local Similarity 97.7%; Pred. No. 3.6e-71;
Matches 297; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
```

```

QY 1661 GCCAGGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCTCC 1720
    |||||||
DB 479 GCCAGGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCTCC 420

QY 1721 TCGTTTATTTTGGTTAACTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAG 1780
    |||||||
DB 419 TCGTTTATTTTGGTTAACTTATGAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAS 360

QY 1781 ATTCAGAGATGGAACACACAGACCCCAAGATCAAAAGCCCAACCATGCCAGCCCTCCC 1840
    |||||||
DB 359 ATTCAGAGATGGAACACACAGACCCCAAGATCAAAAGCCCAACCATGCCAGCCCTCCC 300

QY 1841 AGCACCCCAAGCCCCACGACCATCGTTCGAATTTGACGACACCGTGAGCTTGCTTTG 1900
    |||||||
DB 299 AACACCCCAAGCCCCACGACCATCGTTCGAATTTGACGACACCGTGAGCTTGCTTTG 240
```


Qy 1901 TACTTAACTCATGGAAGATACTACCTTCACGTTTGAATAATGTTTCCGTGGA 1960
Db 239 TACTTTAACTCATGGAAGATACTACCTTCACGTTTGAATAATGTTTCCGTGGA 180
Qy 1961 AATG 1964
Db 179 AATG 176

RESULT 15

US-09-925-065A-198635
; Sequence 198635, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198635
; LENGTH: 469
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-198635

Query Match 15.0%; Score 293.8; DB 6; Length 469;
Best Local Similarity 98.7%; Pred. No. 4.1e-71;
Matches 295; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1666 GGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGCTCCTGTT 1725
Db 1 GGCTCATGGCTGATAATTTCCCTAGGCTTAACAACCAAGCAAGCTTCGCTCCTGTT 60
Qy 1726 TTATTTTGGTTAACTTATGAAAATGTATTAGAAAAGAGTGACGCTCGAGAGATTCA 1785
Db 61 TTATTTTGGTTAACTTATGAAAATGTATTAGAAAAGAGTGACGCTCGAGAGATTCA 120
Qy 1786 GAGATGGAACACACAGACCCAGATCACAAGCCACCATGCCAGCCCTCCAGCAC 1845
Db 121 GAGATGGAACACACAGACCCAGATCACAAGCCACCATGCCAGCCCTCCAGCAC 180
Qy 1846 CCCAGCCCCAGACCATCGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTT 1905
Db 181 CCCAGCCCCAGACCATCGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTT 240
Qy 1906 TAACTCATGGAAGATTAACCTTCACGTTTGAATAATGTTTCCGTGGAATG 1964
Db 241 CAACTCATGGAAGATTAACCACTTCATGTTTGAATAATGTTTCCGTGGAATG 299

Search completed: April 5, 2006, 14:42:32
Job time : 1321 secs

US-10-071-838-1.rnpbn

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:18:49 ; Search time 811 Seconds
(without alignments)
16139.897 Million cell updates/sec

Title: US-10-071-838-1

Perfect score: 1964

Sequence: 1 acggacgtgtagagtcgc.....aaatgttctctgtgaatcg 1964

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1964	100.0	1964	6	ABBS54706
2	1964	100.0	1964	10	ADCC37382
3	1962.4	99.9	1964	10	ADCC37384
4	1948	99.2	2072	5	ABX71403
5	1944	99.0	2064	13	ADQ84596
6	1944	99.0	2064	13	ADQ86595
7	1929.8	98.3	1993	3	AAA15001
8	1929.8	98.3	1993	13	ADT88069
9	1902	96.8	2304	10	ADCC37380
10	1900.4	96.8	2304	10	ADCC37378
11	1836.4	93.5	2084	6	ABZ11810
12	1836.4	93.5	2084	12	ADMA4328
13	1812.4	92.3	2072	6	ABZ11809
14	1812.4	92.3	2072	12	ADMA4327
15	1809	92.1	2647	10	ADCC37388
16	1807.4	92.0	2647	10	ADCC37386
17	1807.4	92.0	2647	11	ADM01991
18	1794	91.3	2146	6	ABBS54707
19	1750	89.1	1862	6	ABBS54708

20	1638.8	83.4	1752	6	ABS78751	AbS78751 DNA encod
21	1220	62.1	7856	5	AAS85531	Aas85531 DNA encod
22	1220	62.1	7856	5	AAS88213	Aas88213 DNA encod
23	1220	62.1	7878	6	ABL64684	AbL64684 Stomach c
24	1218.4	62.0	2881	13	ACN42882	Acn42882 Human dia
25	1128.4	57.5	3799	13	ACN42884	Acn42884 Human dia
26	1127.6	57.4	3901	13	ACN42883	Acn42883 Human dia
27	999	50.9	8180	5	AAS88212	Aas88212 DNA encod
28	999	50.9	8201	2	AAT12170	Aat12170 pJG4-5-CD
29	999	50.9	8201	10	ADE85053	Ade85053 Farnesy1
30	999	50.9	8284	5	AAS85535	Aas85535 DNA encod
31	999	50.9	8408	4	AAL57848	Aal57848 Human pol
32	999	50.9	8420	4	AAL57847	Aal57847 Human pol
33	999	50.9	8435	4	AAL59633	Aal59633 Human pol
34	999	50.9	8435	4	AAL59634	Aal59634 Human pol
35	999	50.9	8435	10	ADCC31883	Adc31883 Human nov
36	882.8	44.9	9804	4	AAS36329	Aas36329 Human car
37	882.8	44.9	9804	4	AAL05173	Aal05173 Human rep
38	882.8	44.9	9804	4	ABL98060	AbL98060 Human tes
39	882.8	44.9	9804	10	ADE47023	Ade47023 Human car
40	882.8	44.9	9804	13	ADJ08441	Adj08441 Human car
41	879.6	44.8	9404	4	AAS36326	Aas36326 Human car
42	879.6	44.8	9404	4	AAL05171	Aal05171 Human rep
43	879.6	44.8	9404	4	ABL98058	AbL98058 Human tes
44	879.6	44.8	9404	10	ADE47020	Ade47020 Human car
45	879.6	44.8	9404	13	ADJ08438	Adj08438 Human car

ALIGNMENTS

RESULT 1	ABS54706	standard; cDNA; 1964 BP.
ID	ABS54706	standard; cDNA; 1964 BP.
AC	ABS54706;	
XX		
DT	28-NOV-2002	(first entry)
XX		
DE	CDNA encoding human PRC17 protein.	
XX		
KW	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;	
KW	breast cancer; cytostatic; chromosome 17q11-12; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	location/Qualifiers
FT	CDS	1..1650
FT		/*tag= a
FT		/product= "Human PRC17 protein"
XX		
PN	WO200262958-A2.	
XX		
PD	15-AUG-2002.	
XX		
PF	08-FEB-2002; 2002WO-US003457.	
XX		
PR	08-FEB-2001; 2001US-0267615P.	
XX		
PA	(TUL-) TULARIK INC.	
XX		
PI	Li J, Powers S, Xiang P, Peng Y;	
XX		
DR	WPI; 2002-706902/76.	
DR	P-PSDB; ABG70736.	
XX		
PT	Novel isolated PRC17 polypeptide useful-diagnostically or prognostically	
PT	to detect diseases or conditions associated with altered PRC17 activity	
PT	or expression relative-to-normal, for example-cancer.	
XX		
PS	Claim 24; Page 62; 78pp; English.	
XX		
CC	The present invention relates to a new PRC17 polypeptide. The invention	

is useful for detecting cancer cells (such as prostate tissue, breast tissue, lung tissue, ovarian tissue) in a biological sample. The invention is further useful for monitoring the efficacy of a therapeutic treatment of a cancer (prostate cancer, ovarian cancer, lung cancer, breast cancer). The molecules of the invention are useful diagnostically or prognostically to detect diseases or conditions associated with altered PRC17 activity or expression relative to normal, for example cancer. The present nucleic acid sequence represents the human PRC17 gene located on chromosome 17q11-12. This sequence encodes the human PRC17 protein of the invention

Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 1964; DB 6; Length 1964;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGTGTAGAGTTCGGCGGACGTTGGTGCCACAAGACGAGACATCATATG 60
DB 1 ATGACGTGTAGAGTTCGGCGGACGTTGGTGCCACAAGACGAGACATCATATG 60
QY 61 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTAAGCCTTTTGA 120
DB 61 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTAAGCCTTTTGA 120
QY 121 AGCTACAACACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTG 180
DB 121 AGCTACAACACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTG 180
QY 181 ACTGCGCGGGAGCGCAAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 240
DB 181 ACTGCGCGGGAGCGCAAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 240
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 300
DB 241 ATGCTGGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 300
QY 301 GGAATGCCCATGAACATCCGGGGGCCGATGTGTCACTCTCTGAACATTGAGGAATG 360
DB 301 GGAATGCCCATGAACATCCGGGGGCCGATGTGTCACTCTCTGAACATTGAGGAATG 360
QY 361 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAAAGGCCAAGAGTCACTGAG 420
DB 361 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAAAGGCCAAGAGTCACTGAG 420
QY 421 CACATCCAGCGCATCGACCGGACGTAACGGGACATTAAAGCAATATATTCTTCAAG 480
DB 421 CACATCCAGCGCATCGACCGGACGTAACGGGACATTAAAGCAATATATTCTTCAAG 480
QY 481 GATCGATACGGAACCAAGCAGCGGAACTAATCCACATCTCTGGCATATGAGGATAT 540
DB 481 GATCGATACGGAACCAAGCAGCGGAACTAATCCACATCTCTGGCATATGAGGATAT 540
QY 541 AACCCGAGGTGGCTACTGACGGACCTGAGCCACATCGCGCTTGTCTCTCTAT 600
DB 541 AACCCGAGGTGGCTACTGACGGACCTGAGCCACATCGCGCTTGTCTCTCTAT 600
QY 601 CTTCTGAGAGAGATGATCTTGGGCACTGGTGCAAGCTGCTGCTGAGAGGCACTCC 660
DB 601 CTTCTGAGAGAGATGATCTTGGGCACTGGTGCAAGCTGCTGCTGAGAGGCACTCC 660
QY 661 CTGCAAGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCAACAGAG 720
DB 661 CTGCAAGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCAACAGAG 720
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGACAAGAAAGATCTATGT 780
DB 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGACAAGAAAGATCTATGT 780
QY 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
DB 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840

QY 841 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAACGCCAACAGCGTTGATGCCGATA 900
DB 841 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAACGCCAACAGCGTTGATGCCGATA 900
QY 901 ACAAGAAATCGCCTTTAAGGTTCAAGAGAGCGGCTTCAAGAACGTCCAGGTGCCCCG 960
DB 901 ACAAGAAATCGCCTTTAAGGTTCAAGAGAGCGGCTTCAAGAACGTCCAGGTGCCCCG 960
QY 961 TGGCAGCTTTTTCGAACCGGTTGTTGATACCTGGGCCAGGATGAGGACATGTGCTC 1020
DB 961 TGGCAGCTTTTTCGAACCGGTTGTTGATACCTGGGCCAGGATGAGGACATGTGCTC 1020
QY 1021 AAGCATTTAGGGCCTTATGAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 1080
DB 1021 AAGCATTTAGGGCCTTATGAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 1080
QY 1081 GCCAACCAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCACTGCGGGAAG 1140
DB 1081 GCCAACCAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCACTGCGGGAAG 1140
QY 1141 ACCCTCTGCAAGGGGACAGGACGCCCTCCAGGCCCAAGCCCGTTCCCGCGGCC 1200
DB 1141 ACCCTCTGCAAGGGGACAGGACGCCCTCCAGGCCCAAGCCCGTTCCCGCGGCC 1200
QY 1201 ATTGTGACCTTCCCGCCAGCGGCACTCGTTCTTCCACAACCTGTCTGTGGGGCT 1260
DB 1201 ATTGTGACCTTCCCGCCAGCGGCACTCGTTCTTCCACAACCTGTCTGTGGGGCT 1260
QY 1261 GTCCGGAGACACCTACCTGTGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320
DB 1261 GTCCGGAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320
QY 1321 GGAGACCTCAGGGTCTCTGAGATTCTCTCAGTGAATCCATGCCCGGCTTCCAAG 1380
DB 1321 GGAGACCTCAGGGTCTCTGAGATTCTCTCAGTGAATCCATGCCCGGCTTCCAAG 1380
QY 1381 GACCTGACGTAGAGGGCCCTTGCTTCCGCAATTAATTCAGACAGAGCTGTGGGTC 1440
DB 1381 GACCTGACGTAGAGGGCCCTTGCTTCCGCAATTAATTCAGACAGAGCTGTGGGTC 1440
QY 1441 CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1500
DB 1441 CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1500
QY 1501 CGGATGATCGGCTTTCGTCGACCCAGCACTGATTCGACAGGCAACCCCTTCAGA 1560
DB 1501 CGGATGATCGGCTTTCGTCGACCCAGCACTGATTCGACAGGCAACCCCTTCAGA 1560
QY 1561 GCTAGGACGAACAGCGTGTGCTCCACCTCAGGCGCTTGCTGCGGCTTCACTTG 1620
DB 1561 GCTAGGACGAACAGCGTGTGCTCCACCTCAGGCGCTTGCTGCGGCTTCACTTG 1620
QY 1621 GAAAGTTCTAGTTCCTCCAGGCTTCTAGAAGCATTCGGGCCAAGGCTCATGGCTGAT 1680
DB 1621 GAAAGTTCTAGTTCCTCCAGGCTTCTAGAAGCATTCGGGCCAAGGCTCATGGCTGAT 1680
QY 1681 AATTCCCTAGGCTTAACAACCAAGCAAGCTTGGCGTCTGTTTATTTTGGTTAAA 1740
DB 1681 AATTCCCTAGGCTTAACAACCAAGCAAGCTTGGCGTCTGTTTATTTTGGTTAAA 1740
QY 1741 CTTATGAAATGTATTAAAGAAAGTGACGCTCGAAGAGATTCAAGATGAAACACAC 1800
DB 1741 CTTATGAAATGTATTAAAGAAAGTGACGCTCGAAGAGATTCAAGATGAAACACAC 1800
QY 1801 AGACCCGATCAACAAGCCACATGCCAGGCCCTCCAGCAACCCCAAGCCCAAGAC 1860
DB 1801 AGACCCGATCAACAAGCCACATGCCAGGCCCTCCAGCAACCCCAAGCCCAAGAC 1860
QY 1861 CATGTTCTGAATTTCTGACGACACCGTGAAGCTGCTTGTACTTTAACTCATGGAAG 1920
DB 1861 CATGTTCTGAATTTCTGACGACACCGTGAAGCTGCTTGTACTTTAACTCATGGAAG 1920
QY 1921 ATAACTACTTCAAGTTTGAATAAATGTTTCTGTGGAATG 1964

Db 1921 ATAACTACTTCACGTTTGAATAATGTTCTCTGTGAATG 1964

RESULT 2
ADC37382
ID ADC37382 standard; DNA; 1964 BP.

XX ADC37382;
AC
XX 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 215.

KM Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KM cancer; infectious disease; bone disease; AIDS;
KM neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KM Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KM Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

XX Homo sapiens.

XX WO2003048202-A2.

XX 12-JUN-2003.

PF 03-DEC-2002; 2002MO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

XX (ASAH) ASAMI KASEI KK.

PI Matsuda A, Muramatsu S;

DR WPI; 2003-505282/47.

PT P-PSDB; ADC37383.

PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.

PS Claim 4; SEQ ID NO 215; 938pp; English.

CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.

XX Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;

Query Match 100.0%; Score 1964; DB 10; Length 1964;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACGTGCTAGAGTCCGCGGACAGTGTGGGCAACAAGCGAGAGACATCATATG 60
DB 1 ATGACGTGCTAGAGTCCGCGGACAGTGTGGGCAACAAGCGAGAGACATCATATG 60
QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 120
DB 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 120
QY 121 AGCTAACAAACAAGCTGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
DB 121 AGCTAACAAACAAGCTGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
QY 181 ACTGCGCGGAGCGAAGCAATTCCGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 240

Db 181 ACTGCGCGGAGCGAAGCAATTCCGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 240
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAAGAAAGCTCATAGTCGAGCGTACAAG 300
Db 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAAGAAAGCTCATAGTCGAGCGTACAAG 300
QY 301 GGAATGCCCATGAACATCCGGGCGCGATGTGTGTCAGTCTCTGAAACATTGAGGAATG 360
Db 301 GGAATGCCCATGAACATCCGGGCGCGATGTGTGTCAGTCTCTGAAACATTGAGGAATG 360
QY 361 AAGTTGAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 420
Db 361 AAGTTGAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 420
QY 421 CACATCCAGCCCATCGACCCGGGACGTAAAGCGGACATTAAGGAAGCATATATTTCAAG 480
Db 421 CACATCCAGCCCATCGACCCGGGACGTAAAGCGGACATTAAGGAAGCATATATTTCAAG 480
QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTGTCATATGAGAGTAT 540
Db 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTGTCATATGAGAGTAT 540
QY 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
QY 601 CTTCCTGAGAGATGATCTTGGGCACTGTGTCAGCTGTGGCCAGTGAAGGACTCC 660
Db 601 CTTCCTGAGAGATGATCTTGGGCACTGTGTCAGCTGTGGCCAGTGAAGGACTCC 660
QY 661 CTGCAGGATTTTCAAGCCCAATGGCGGACCGTCCAGAGGGCTCCAAGACCAAGAG 720
Db 661 CTGCAGGATTTTCAAGCCCAATGGCGGACCGTCCAGAGGGCTCCAAGACCAAGAG 720
QY 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGACCAAGAAAGTCTATGT 780
Db 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGACCAAGAAAGTCTATGT 780
QY 781 GGGCAGTGTTCCTCCGTTAAGGCTGCTCATCCGGATATGATTGACGGGATCTCTCGGG 840
Db 781 GGGCAGTGTTCCTCCGTTAAGGCTGCTCATCCGGATATGATTGACGGGATCTCTCGGG 840
QY 841 CTCACCTGCGCCTGTGGGACGTGTATCTGTAGAAAGCGAAGCGGTTGATGCCGATA 900
Db 841 CTCACCTGCGCCTGTGGGACGTGTATCTGTAGAAAGCGAAGCGGTTGATGCCGATA 900
QY 901 ACAAGATCGCCTTTAAGTTACAGCAGAGCGCCTTACGAAAGCTCCAGTGTGCCCG 960
Db 901 ACAAGATCGCCTTTAAGTTACAGCAGAGCGCCTTACGAAAGCTCCAGTGTGCCCG 960
QY 961 TGGGCAAGTTTTCACACCGGTTCTGTGATACCTGGGCAAGGATGAGACATGTGCTC 1020
Db 961 TGGGCAAGTTTTCACACCGGTTCTGTGATACCTGGGCAAGGATGAGACATGTGCTC 1020
QY 1021 AAGCATTTAGGGCTTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1021 AAGCATTTAGGGCTTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
QY 1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTACGTGGCGGAAG 1140
Db 1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTACGTGGCGGAAG 1140
QY 1141 ACCCTCTGCAAGGGGACAGGACCCCTCCAGGCCCAACAGCCCGTTCCCGGGCCC 1200
Db 1141 ACCCTCTGCAAGGGGACAGGACCCCTCCAGGCCCAACAGCCCGTTCCCGGGCCC 1200
QY 1201 ATTGTCAGCTTCCCGCCACGGGACCTGTTCTTCCACACCTGTCTGTGGGCT 1260
Db 1201 ATTGTCAGCTTCCCGCCACGGGACCTGTTCTTCCACACCTGTCTGTGGGCT 1260
QY 1261 GTCCGGGAAGACACTTACCTGTGGGACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320

Db 1261 GTCCGGAGAGACCTACCCCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
QY 1321 GGAGGACCTCAGGGTCTCTGAGATTCTGAGTGAATCCATGCCCCGCTCCCAACG 1380
Db 1321 GGAGGACCTCAGGGTCTCTGAGATTCTGAGTGAATCCATGCCCCGCTCCCAACG 1380
QY 1381 GACCTGGACGTAGAGGGCCCTTGGTCCGCCATTATGATTTTCAGACAGAGCTGTGGGTC 1440
Db 1381 GACCTGGACGTAGAGGGCCCTTGGTCCGCCATTATGATTTTCAGACAGAGCTGTGGGTC 1440
QY 1441 CGTGCCATATCCAGAGAGACCACTGCCCCCTGCTGGCAGGCTGAACACCCCTGGGAG 1500
Db 1441 CGTGCCATATCCAGAGAGACCACTGCCCCCTGCTGGCAGGCTGAACACCCCTGGGAG 1500
QY 1501 CGGGTGAATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1560
Db 1501 CGGGTGAATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA 1560
QY 1561 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCCCTTGCGGCTCCACTTG 1620
Db 1561 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCCCTTGCGGCTCCACTTG 1620
QY 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
Db 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
QY 1681 AATTTCCCTAGGCTTAAACAACCCAGCAAGCTTCGCTCCTGTTTATTTTGTAA 1740
Db 1681 AATTTCCCTAGGCTTAAACAACCCAGCAAGCTTCGCTCCTGTTTATTTTGTAA 1740
QY 1741 CTTATGAATAATGTTAAGAAAGTGCAGCTCGAGAGATTTCAGAGATGGAACACACC 1800
Db 1741 CTTATGAATAATGTTAAGAAAGTGCAGCTCGAGAGATTTCAGAGATGGAACACACC 1800
QY 1801 AGACCCCAATCAAAAGCCAAACCATGCCCCAGCCCCCTCCAGACACCCCAAGCAGAC 1860
Db 1801 AGACCCCAATCAAAAGCCAAACCATGCCCCAGCCCCCTCCAGACACCCCAAGCAGAC 1860
QY 1861 CATCGTTCGAATTCGACGACACCGTGAGCCTGCTTGTACTTTAAACTCATGGAAG 1920
Db 1861 CATCGTTCGAATTCGACGACACCGTGAGCCTGCTTGTACTTTAAACTCATGGAAG 1920
QY 1921 ATAACCTACCTCAGCTTTTGAATAATGTTTCTGTTGAATG 1964
Db 1921 ATAACCTACCTCAGCTTTTGAATAATGTTTCTGTTGAATG 1964

RESULT 3
AD3C37384 standard; DNA; 1964 BP.
XX AD3C37384;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nucleic acid factor kappa B (NF-kappaB) activating gene, SEQ ID 217.
XX
XX Nucleic acid factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiac; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PSDB; AD3C37385.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 217; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (AD3C37168-AD3C37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
XX
SQ Sequence 1964 BP; 473 A; 552 C; 561 G; 378 T; 0 U; 0 Other;

Query Match 99.9%; Score 1962.4; DB 10; Length 1964;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1963; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGCTGTGTAGAGGTCCGGGCAAGTTGGTGGCAAGAGGAGACATCATATG 60
Db 1 ATGAGCTGTGTAGAGGTCCGGGCAAGTTGGTGGCAAGAGGAGACATCATATG 60
QY 61 AATATGAAAAGGAGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 120
Db 61 AATATGAAAAGGAGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTGA 120
QY 121 AGCTACAAACAACGTCATCTTTGGGATTTGATGATGAGACGAGCTGCTCTG 180
Db 121 AGCTACAAACAACGTCATCTTTGGGATTTGATGATGAGACGAGCTGCTCTG 180
QY 181 ACTGCGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240
Db 181 ACTGCGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240
QY 241 ATGCTGGGAGCTGGAGAAATACAAAGCAGACGAAAGCTCATAGATCGACGTACAAG 300
Db 241 ATGCTGGGAGCTGGAGAAATACAAAGCAGACGAAAGCTCATAGATCGACGTACAAG 300
QY 301 GGAATGCCATGAACATCCGGGCCCCGATGTGTGTCATCTCTGAACATGAGGAATG 360
Db 301 GGAATGCCATGAACATCCGGGCCCCGATGTGTGTCATCTCTGAACATGAGGAATG 360
QY 361 AAGTTGAAAAACCCCGAAGATACCATCATGAAAGAGGCAAGAGTCACTGAG 420
Db 361 AAGTTGAAAAACCCCGAAGATACCATCATGAAAGAGGCAAGAGTCACTGAG 420
QY 421 CACATCAGCGCATCGACCGGAGCTTAAGCGGAGATTAAGAGCATATATCTTCAGG 480
Db 421 CACATCAGCGCATCGACCGGAGCTTAAGCGGAGATTAAGAGCATATATCTTCAGG 480
QY 481 GATGATACGGAACCAAGCAGCGGAATCTCAATCTCTGTCATATGAGAGTAT 540
Db 481 GATGATACGGAACCAAGCAGCGGAATCTCAATCTCTGTCATATGAGAGTAT 540
QY 541 AACCCGAGGTGGCTACTGCAAGGACCTGAGCAATCGCCGCTTCTCTCTAT 600
Db 541 AACCCGAGGTGGCTACTGCAAGGACCTGAGCAATCGCCGCTTCTCTCTAT 600
QY 601 CTTCTGAGGAGATGATCTTGGGCACTGGTGACGCTGGCCAGTGAAGGCACTCC 660
Db 601 CTTCTGAGGAGATGATCTTGGGCACTGGTGACGCTGGCCAGTGAAGGCACTCC 660

QY 661 CTGAGGATTTTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
 DB 661 CTGAGGATTTTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
 QY 721 CATGTGTAAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 780
 DB 721 CATGTGTAAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 780
 QY 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATGATGACGGATCTCTCGGG 840
 DB 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATGATGACGGATCTCTCGGG 840
 QY 841 CTCACTGCTGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAAGCGGCTGATGCCGATA 900
 DB 841 CTCACTGCTGCGCTGTGGAGCGTGTATCTGTAGAAAGCGAAGCGGCTGATGCCGATA 900
 QY 901 ACAAGATCGCCTTTAAGTTCAAGAGCGGCTCAAGAGACGTCCAGGTGGCCCG 960
 DB 901 ACAAGATCGCCTTTAAGTTCAAGAGCGGCTCAAGAGACGTCCAGGTGGCCCG 960
 QY 961 TGGGACGTTTTTTCACACCGGTTCTGTATACCTGGGCGAGGATGAGACACTGTCTC 1020
 DB 961 TGGGACGTTTTTTCACACCGGTTCTGTATACCTGGGCGAGGATGAGACACTGTCTC 1020
 QY 1021 AAGCATCTTAGGGCTCTATGAGAAATTAACAAGAAAGAGGGGACCTGCCAATCCCA 1080
 DB 1021 AAGCATCTTAGGGCTCTATGAGAAATTAACAAGAAAGAGGGGACCTGCCAATCCCA 1080
 QY 1081 GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCGGCTTAAGTGGCGGAAG 1140
 DB 1081 GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCGGCTTAAGTGGCGGAAG 1140
 QY 1141 ACCCTCTGCAAGGGGAGCAAGGAGGCGGCTCCAGGCGCAACAGCCGGTTCCCGGGCCC 1200
 DB 1141 ACCCTCTGCAAGGGGAGCAAGGAGGCGGCTCCAGGCGCAACAGCCGGTTCCCGGGCCC 1200
 QY 1201 ATTGTCAGCTTCCCGGCAAGGGGACCTGTTCTTCCACAACCTGTCTGTGGGGCT 1260
 DB 1201 ATTGTCAGCTTCCCGGCAAGGGGACCTGTTCTTCCACAACCTGTCTGTGGGGCT 1260
 QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGGCTCAG 1320
 DB 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGGCTCAG 1320
 QY 1321 GAGGACCTCAGGGTCTGTGAGATTCCTGACGTGAATCCATGCCCCGCTCCCAACG 1380
 DB 1321 GAGGACCTCAGGGTCTGTGAGATTCCTGACGTGAATCCATGCCCCGCTCCCAACG 1380
 QY 1381 GACCTGACGTAGAGGGGCTTGTGCTCCGCATTAATTCAGACAGAGCTGTGGTTC 1440
 DB 1381 GACCTGACGTAGAGGGGCTTGTGCTCCGCATTAATTCAGACAGAGCTGTGGTTC 1440
 QY 1441 CGTCCCATATCCAGAGAGACAGCTGGCCCCCTGTGCGAGGCTGAACACCTGCGGAG 1500
 DB 1441 CGTCCCATATCCAGAGAGACAGCTGGCCCCCTGTGCGAGGCTGAACACCTGCGGAG 1500
 QY 1501 CGGGTGAATCGGCTTTCGCTGCAACCCAGCACTGATTCGACAGGGGACCCCTTCAGA 1560
 DB 1501 CGGGTGAATCGGCTTTCGCTGCAACCCAGCACTGATTCGACAGGGGACCCCTTCAGA 1560
 QY 1561 GCTAGGAGCAAGACAGCGGTGTGCTCCACCTCAGGCGCTTGCCTGTGCGGCTCCACTTG 1620
 DB 1561 GCTAGGAGCAAGACAGCGGTGTGCTCCACCTCAGGCGCTTGCCTGTGCGGCTCCACTTG 1620
 QY 1621 GAAAGTTCTCAGTTCCCTCAGAGCTTCTAGAAAGATCTGGGCGAGGCTCATGGCTGAT 1680
 DB 1621 GAAAGTTCTCAGTTCCCTCAGAGCTTCTAGAAAGATCTGGGCGAGGCTCATGGCTGAT 1680
 QY 1681 AATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCCGCTCTCTGTTTATTTTGGTTAA 1740
 DB 1681 AATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCCGCTCTCTGTTTATTTTGGTTAA 1740

QY 1741 CTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAGATTTCAGAGATGGAACACACC 1800
 DB 1741 CTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGAGATTTCAGAGATGGAACACACC 1800
 QY 1801 AGACCCAGATCAACAAGCCATATGCCAGCCCTCCAGACACCCCCAGCCCCACGAC 1860
 DB 1801 AGACCCAGATCAACAAGCCATATGCCAGCCCTCCAGACACCCCCAGCCCCACGAC 1860
 QY 1861 CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG 1920
 DB 1861 CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG 1920
 QY 1921 ATAACTACCTTCACGTTTGAATAATGTTTCTGTGAATG 1964
 DB 1921 ATAACTACCTTCACGTTTGAATAATGTTTCTGTGAATG 1964

RESULT 4
 ABX71403
 ID ABX71403 standard; cDNA; 2072 BP.
 XX
 AC ABX71403;
 XX
 DT 14-APR-2003 (first entry)
 XX
 DE Human cell cycle-associated cDNA from clone DKFzphes3_35p22.
 XX
 KW Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200112659-A2.
 XX
 PD 22-FEB-2001.
 XX
 PF 18-AUG-2000; 2000WO-IB001496.
 XX
 PR 18-AUG-1999; 99US-0149499P.
 XX
 PR 28-SEP-1999; 99US-0156503P.
 XX
 PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
 XX
 PI Wiemann S;
 XX
 DR MPI; 2001-327840/34.
 DR P-PSDB; ABUS3234.
 XX
 PT Nucleic acids having the sequences of clones isolated from libraries of
 PT different human tissues, useful in recombinant DNA methodologies.
 XX
 PS Claim 1; Page 867; 1095pp; English.
 XX
 CC This invention describes novel polynucleotides and polypeptides isolated
 CC from human cDNA libraries which can be used for gene therapy or in
 CC vaccines. The polynucleotides of the invention and antibodies encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The products of the
 CC invention may also be used to identify modulators of expression and
 CC activity and to down regulate expression and activity. The antibodies of
 CC the invention may also be used as diagnostic agents for detecting the
 CC presence of polypeptides in samples. This sequence encodes a polypeptide
 CC described in the disclosure of the invention
 XX
 SQ Sequence 2072 BP; 510 A; 579 C; 586 G; 397 T; 0 U; 0 Other;

Query Match 99.2%; Score 1948; DB 5; Length 2072;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1954; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGACGCTGTAGAGTCCGGGCAAGTTGTGGGCAACAAGCGAGAGACATCATATG 60
 DB 99 ATGACGCTGTAGAGTCCGGGCAAGTTGTGGGCAACAAGCGAGAGACATCATATG 158

Qy 61 AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCCTAAGCCTTTTCGA 120
|
Db 159 AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCCTAAGCCTTTTCGA 218
|
Qy 121 AGCTACAAACAACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
|
Db 219 AGCTACAAACAACGTCGATCATTTGGGATTGTACATGAGACGAGCTGCTCTCTG 278
|
Qy 181 ACTGCGCGGAGCGCAAGCAAAATTCGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 240
|
Db 279 ACTGCGCGGAGCGCAAGCAAAATTCGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 338
|
Qy 241 ATGCTGGAGACTGGGAGAAATACAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 300
|
Db 339 ATGCTGGAGACTGGGAGAAATACAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 398
|
Qy 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTGAAACATTGAGGAATG 360
|
Db 399 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTGAAACACTGAGGAATG 458
|
Qy 361 AAGTTGAAAAACCCCGAAGATACCAATCATGAAGAGAAAGGCAAGGTCATCTGAG 420
|
Db 459 AAGTTGAAAAACCCCGAAGATACCAATCATGAAGAGAAAGGCAAGGTCATCTGAG 518
|
Qy 421 CACATCCAGCGCATCGACCGGACGTAAACGGGACATTAAAGAAACATATATTCTTCAAG 480
|
Db 519 CACATCCAGCGCATCGACCGGACGTAAACGGGACATTAAAGAAACATATATTCTTCAAG 578
|
Qy 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGATAT 540
|
Db 579 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGATAT 638
|
Qy 541 AACC CGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
|
Db 639 AACC CGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 698
|
Qy 601 CTTCTGAGAGGATGCAATCTGGGCACTGTGCAAGCTGCTGGCCAGTGAGAGGCACTCC 660
|
Db 699 CTTCTGAGAGGATGCAATCTGGGCACTGTGCAAGCTGCTGGCCAGTGAGAGGCACTCC 758
|
Qy 661 CTGCAAGGATTTCACAGCCCAAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 720
|
Db 759 CTGCAAGGATTTCACAGCCCAAATGCGGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 818
|
Qy 721 CATGTGTAGCCACGTCAAAACCAAGACCATGGGCAATGAGCAAGAAAGATCTATGT 780
|
Db 819 CATGTGTAGCCACGTCAAAACCAAGACCATGGGCAATGAGCAAGAAAGATCTATGT 878
|
Qy 781 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGAATTGACGGGATCTCTCGGG 840
|
Db 879 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGAATTGACGGGATCTCTCGGG 938
|
Qy 841 CTCACCCCTGGCCCTGTGGGACGTGTATCTGTGAGAAAGGCAAGCGGTTGATGGCGATA 900
|
Db 939 CTCACCCCTGGCCCTGTGGGACGTGTATCTGTGAGAAAGGCAAGCGGTTGATGGCGATA 998
|
Qy 901 ACAAGAAATCGCCTTTAAGGTTACAGAGAAAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 960
|
Db 999 ACAAGAAATCGCCTTTAAGGTTACAGAGAAAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 1058
|
Qy 961 TGGGACGTTTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1020
|
Db 1059 TGGGACGTTTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1118
|
Qy 1021 AAGCATCTTAAGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCCA 1080
|
Db 1119 AAGCATCTTAAGGCTCTATGAAGAACTAACAGAAAGAGGGGGACCTGCCACCCCCA 1178
|
Qy 1081 GCCAAACCCGAGCAAGGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAAAG 1140
|
Db 1179 GCCAAACCCGAGCAAGGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAAAG 1238
|
Qy 1141 ACCCTCTGCAAGGGGAGCAGGCGCCCTCCAGGCCCAAGCCCGGTTCCCGGGGCC 1200

Db 1239 ACCCTCTGCAAGGGGAGCAGGCAAGCCCTCCAGGCCCAACAGCCCGGTTCCCGGGCCCC 1298
|
Qy 1201 ATTTGGTCAGCTTCCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGCT 1260
|
Db 1299 ATTTGGTCAGCTTCCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGCT 1358
|
Qy 1261 GTCCGGAAAGACACTACCTCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTTGCTCAG 1320
|
Db 1359 GTCCGGAAAGACACTACCTCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTTGCTCAG 1418
|
Qy 1321 GGAGACCTCAGGGTCTTGAGATTCTCTGACGTGAACTCCATGCCCGGCTCCCAAG 1380
|
Db 1419 GGAGACCTCAGGGTCTTGAGATTCTCTGACGTGAACTCCATGCCCGGCTCCCAAG 1478
|
Qy 1381 GACCTGACGTAGAGGCGCTTGGTTCGCCCATTAATTAATTCAGACAGAGCTGCTGGGTC 1440
|
Db 1479 GACCTGACGTAGAGGCGCTTGGTTCGCCCATTAATTAATTCAGACAGAGCTGCTGGGTC 1538
|
Qy 1441 CGTGCCATATCCAGAGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGAG 1500
|
Db 1539 CGTGCCATATCCAGAGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGAG 1598
|
Qy 1501 CGGCTGAGATCGGCTTTCGCTGCACCCAGCAGCTGATTCGACCAAGGCAACCCCTTCA 1560
|
Db 1599 CGGCTGAGATCGGCTTTCGCTGCACCCAGCAGCTGATTCGACCAAGGCAACCCCTTCA 1658
|
Qy 1561 GCTAGGACGAACAGCCGCTGTCTCCACCTCAGGCGCTTGCTCTGCGGCTTCACTTG 1620
|
Db 1659 GCTAGGACGAACAGCCGCTGTCTCCACCTCAGGCGCTTGCTCTGCGGCTTCACTTG 1718
|
Qy 1621 GAAAGTCTCAGTTCCTCTCAGGCTTCTAAGACATCTGGCCAGGCGCTCATGGCTGAT 1680
|
Db 1719 GAAAGTCTCAGTTCCTCTCAGGCTTCTAAGACATCTGGCCAGGCGCTCATGGCTGAT 1778
|
Qy 1681 AATTTCCTTAGGCTTAAACAACCAAGCTTGCCTCTGTTTATTTTGGTTAA 1740
|
Db 1779 AATTTCCTTAGGCTTAAACAACCAAGCTTGCCTCTGTTTATTTTGGTTAA 1838
|
Qy 1741 CTTATGAATAATGATTAAAGAAAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
|
Db 1839 CTTATGAATAATGATTAAAGAAAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACACC 1898
|
Qy 1801 AGACCCCAAGATCAAAAGCCCAACCATGCCAGGCCCTCCAGCAACCCCAAGGAC 1860
|
Db 1899 AGACCCCAAGATCAAAAGCCCAACCATGCCAGGCCCTCCAGCAACCCCAAGGAC 1958
|
Qy 1861 CATGTTCTGAATTCTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
|
Db 1959 CATGTTCTGAATTCTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 2018
|
Qy 1921 ATTAATACTTCAAGTTTGAATAAATGTTTCTGTGAATG 1964
|
Db 2019 ATTAATACTTCAAGTTTGAATAAATGTTTCTGTGAATG 2062
|
RESULT 5
ADQ84596
ID ADQ84596 standard; cDNA; 2064 BP.
XX
AC ADQ84596;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1410.
XX
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX

PD 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX Wu TD, Zhou Y;
PI WPI; 2004-534300/51.
XX
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 1410; 5504pp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;

Query Match 99.0%; Score 1944; DB 13; Length 2064;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1959; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 ATGAGCTGTGTAGAGTCCGGGAGTTGTGGGCACAGAAGCGAGAGACATCATTTATG 60
DB 99 ATGAGCTGTGTAGAGTCCGGGAGTTGTGGGCACAGAAGCGAGAGACATCATTTATG 158

QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGACAAGGGGCTTAAGCCTTTTGA 120
DB 159 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGACAAGGGGCTTAAGCCTTTTGA 218

QY 121 AGCTACAACAACAGCTGATCATTTGGGATGTGACA--TGAGACGAGCTGCTCTC 178

DB 219 AGCTACAACAACAACGTCGATCATTTGGGATTTGTACAAGTGAACGAGCTGCTCTC 278

QY 179 TGACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGG 238
DB 279 TGACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGG 338

QY 239 ATATGCTGGGAGACTGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGCTACA 298
DB 339 ATATGCTGGGAGACTGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGCTACA 398

QY 299 AGGGAATGCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGACATTTGAGAAA 358
DB 399 AGGGAATGCCATGAACATCCGGGGCCCGATGTGTCACTCTCTGACATTTGAGAAA 458

QY 359 TGAAGTTGAAAAAACCAGGAATACCATGATCATGAAGAGAAAGGCAAGGTCATCTG 418
DB 459 TGAAGTTGAAAAAACCAGGAATACCATGATCATGAAGAGAAAGGCAAGGTCATCTG 518

QY 419 AGCACATCCAGCCGATCGACCCGGAGCTAAGCGGGACATTAAGGAACATATATCTTCA 478
DB 519 AGCACATCCAGCCGATCGACCCGGAGCTAAGCGGGACATTAAGGAACATATATCTTCA 578

QY 479 GGGATCGATACGAACCAAGCAGCGGAATCTACTCCACATCTCTGGCATATGAGAGT 538
DB 579 GGGATCGATACGAACCAAGCAGCGGAATCTACTCCACATCTCTGGCATATGAGAGT 638

QY 539 ATAACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTCTCTCTCT 598
DB 639 ATAACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTCTCTCTCT 698

QY 599 ATCTTCTGAGAGAGATGATCTTGGGCACTGTGTGACGTCTGCGCAGTGAGAGCACT 658
DB 699 ATCTTCTGAGAGAGATGATCTTGGGCACTGTGTGACGTCTGCGCAGTGAGAGCACT 758

QY 659 CCCTGCAAGGATTTTACAGAGCCCAATGGCGGACCGTTCAGGGGCTCCAAGACCAACAG 718
DB 759 CCCTGCAAGGATTTTACAGAGCCCAATGGCGGACCGTTCAGGGGCTCCAAGACCAACAG 818

QY 719 AGCATGTGTAGCCACGTCACAAACCAAGACCATGGGCAATGAGCAAGAAATCTAT 778
DB 819 AGCATGTGTAGCCACGTCACAAACCAAGACCATGGGCAATGAGCAAGAAATCTAT 878

QY 779 GTGGGAGTGTCCCGGTTAGGCTGCTCATCCGATATTTGATGACGGGATCTCTCG 838
DB 879 GTGGGAGTGTCCCGGTTAGGCTGCTCATCCGATATTTGATGACGGGATCTCTCG 938

QY 839 GGCTCACCTTGGCGCTGTGGGAGTGTATCTGTGTAGAAAGCGGAACGGCTGTATGCCGA 898
DB 939 GGCTCACCTTGGCGCTGTGGGAGTGTATCTGTGTAGAAAGCGGAACGGCTGTATGCCGA 998

QY 899 TAACAGAATCGCTTTAAGTTTACAGAGAAGCGCTTCAGAAAGCTTCAGGTGTGCC 958
DB 999 TAACAGAATCGCTTTAAGTTTACAGAGAAGCGCTTCAGAAAGCTTCAGGTGTGCC 1058

QY 959 CGTGGCAGCTTTTGTCAACCGGTTCTGTGATATCTGGGCCAGGGATGAGACACTGTGC 1018
DB 1059 CGTGGCAGCTTTTGTCAACCGGTTCTGTGATATCTGGGCCAGGGATGAGACACTGTGC 1118

QY 1019 TCAAGCATCTTAGGCTTATGAAGAACTAAAGAAAGCAGGGGACCTGCCACCCC 1078
DB 1119 TCAAGCATCTTAGGCTTATGAAGAACTAAAGAAAGCAGGGGACCTGCCACCCC 1178

QY 1079 CAGCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCACTGTGGCGGA 1138
DB 1179 CAGCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTCACTGTGGCGGA 1238

QY 1139 AGACCTCTGCAAGGGGAGCAGGCAAGGCCCTTCCAGGCCCAACAGCCGGTTCCCGGGC 1198
DB 1239 AGACCTCTGCAAGGGGAGCAGGCAAGGCCCTTCCAGGCCCAACAGCCGGTTCCCGGGC 1298

QY 1199 CCAATTGCTCAGCTTCCCGGCCACGGGACCTGTTCTTCCACACCTGTCTGTGGGG 1258

Db 1299 CCATTGGTCAGCTTCCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGG 1358
Qy 1259 CTGTCCGGAGACACCTACCCCTGTGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTC 1318
Db 1359 CTGTCCGGAGACACCTACCCCTGTGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTC 1418
Qy 1319 AGGAGGACCTCAGGGTCTCTGAGATTCTGAGTGAATCCATGCCCCGCTGCCAA 1378
Db 1419 AGGAGGACCTCAGGGTCTCTGAGATTCTGAGTGAATCCATGCCCCGCTGCCAA 1478
Qy 1379 CGGACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGG 1438
Db 1479 CGGACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAGACAGAGCTGTGG 1538
Qy 1439 TCCGTGCCATATCCACGAGAGACCAGCTGGCCCCCTGTGGCAGGCTGAACACCCTGCGG 1498
Db 1539 TCCGTGCCATATCCACGAGAGACCAGCTGGCCCCCTGTGGCAGGCTGAACACCCTGCGG 1598
Qy 1499 AGCGGTGAGATCGGCTTTCGTGACCCAGCACTGATTCGACAGGGCACCCTCTCA 1558
Db 1599 AGCGGTGAGATCGGCTTTCGTGACCCAGCACTGATTCGACAGGGCACCCTCTCA 1658
Qy 1559 GAGCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGCGCTTGCCTGTGCGGCTCACT 1618
Db 1659 GAGCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGCGCTTGCCTGTGCGGCTCACT 1718
Qy 1619 TGGAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGG 1678
Db 1719 TGGAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGG 1778
Qy 1679 ATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCGTCTGTTTATTTTGGTTA 1738
Db 1779 ATAATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCAATCTCTGTTTATTTGGTTA 1838
Qy 1739 AACTTATGAAATGTATTAGAAGAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACA 1798
Db 1839 AACTTATGAAATGTATTAGAAGAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACA 1898
Qy 1799 CCAGACCCCAAGATCAAAAGCCCAATGCCAGCCCTCCAGACACCCCGGACCG 1858
Db 1899 CCAGACCCCAAGATCAAAAGCCCAATGCCAGCCCTCCAGACACCCCGGACCG 1958
Qy 1859 ACCATCGTCTGAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAA 1918
Db 1959 ACCATCGTCTGAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTCAACTCATGGAA 2018
Qy 1919 GGATAACTACTCTCACGTTTGAATAATGTTTCTCTGTTGAATG 1964
Db 2019 GGATAACTACTCTCATGTTTGAATAATGTTTCTCTGTTGAATG 2064

RESULT 6
ADQ86595
ID ADQ86595 standard; cDNA; 2064 BP.

XX
AC ADQ86595;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #3468.
XX
KW human; tumour-associated antigenic target; TAT; cyostatic; gene therapy;
XX cancer; cell proliferative disorder; gene; ss.
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.

XX
PA (GETH) GENENTECH INC.
PA (WUTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PS preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 3468; 5504pp; English.

CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that expresses the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cyostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX
SQ Sequence 2064 BP; 498 A; 579 C; 588 G; 399 T; 0 U; 0 Other;

Query Match 99.0%; Score 1944; DB 13; Length 2064;
Best local Similarity 99.6%; Pred. No. 0;
Matches 1959; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 1 ATGAGCGGTGTAAGGTCGGGCAAGTTGGTGGCACAAGAGGAGACATCATATG 60
Db 99 ATGAGCGGTGTAAGGTCGGGCAAGTTGGTGGCACAAGAGGAGACATCATATG 158
Qy 61 AATATGAAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 120
Db 159 AATATGAAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 218
Qy 121 AGTACACAACAACGTCGATCATTTGGGATTGTACA--TGAGACGAGCTCCTCCTC 178
Db 219 AGTACACAACAACGTCGATCATTTGGGATTGTACAAGTGAACGAGCTGCTCCTC 278
Qy 179 TGACTGGCGGGAGGCGAAGCAATTGCGCGGAGATCAGCCGAAAGAGAGTGGTGG 238

Db 279 TGACTGCGCGGAGCGGAAGCAAAATTCGGCGGAGATCACCGAAAGAGCAAGTGGGTGG 338
Qy 239 ATATGCTGGGAGACTGGAGAAATACAAAAGCAGCAAGAAAGCTCATAGATCGAGCGTACA 298
Db 339 ATATGCTGGGAGACTGGAGAAATACAAAAGCAGCAAGAAAGCTCATAGATCGAGCGTACA 398
Qy 299 AGGGAATGCCCATGACATCCGGGGCCCGATGTGTCACTCTCTGAACTTGAAGAAA 358
Db 399 AGGGAATGCCCATGACATCCGGGGCCCGATGTGTCACTCTCTGAACTTGAAGAAA 458
Qy 359 TGAAGTTGAAAAAACCCTGGAAAGATAACAGATCATGAAGAGAAAGGCAAGAGTCACTG 418
Db 459 TGAAGTTGAAAAAACCCTGGAAAGATAACAGATCATGAAGAGAAAGGCAAGAGTCACTG 518
Qy 419 AGCACAATCCAGCGCATCGACCGGGAGCTAAGCGGGACATTAAAGAAACATAATTCTTCA 478
Db 519 AGCACAATCCAGCGCATCGACCGGGAGCTAAGCGGGACATTAAAGAAACATAATTCTTCA 578
Qy 479 GGGATCGATACGGAACCAAGCAGCGGGAACCTACTCCATCTCTGACATATGAGAGT 538
Db 579 GGGATCGATACGGAACCAAGCAGCGGGAACCTACTCCATCTCTGACATATGAGAGT 638
Qy 539 ATAAACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCT 598
Db 639 ATAAACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCT 698
Qy 599 ATCTTCTGAGGAGATGCATTTCTGGGCACTGGTGCACTGTGGCCAGTGAAGGCACT 658
Db 699 ATCTTCTGAGGAGATGCATTTCTGGGCACTGGTGCACTGTGGCCAGTGAAGGCACT 758
Qy 659 CCCTGCAGGGATTTACACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGCAACAGG 718
Db 759 CCCTGCAGGGATTTACACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGCAACAGG 818
Qy 719 AGCATGTGTTAGCCACGTCACAACCCAAAGACCATGGGGCATCAGACAAAGAAATCTAT 778
Db 819 AGCATGTGTTAGCCACGTCACAACCCAAAGACCATGGGGCATCAGACAAAGAAATCTAT 878
Qy 779 GTGGGCAGTGTTCCCGTTAAGGCTGCTCATCCGGAATAATTGACGGGATCTCTCTCG 838
Db 879 GTGGGCAGTGTTCCCGTTAAGGCTGCTCATCCGGAATAATTGACGGGATCTCTCTCG 938
Qy 839 GGCTCACCCCTGCGCTGTGGGACGTGTATCTGTGTAAGGGCAACAGGCGTTGATGCCGA 898
Db 939 GGCTCACCCCTGCGCTGTGGGACGTGTATCTGTGTAAGGGCAACAGGCGTTGATGCCGA 998
Qy 899 TAACAAGAATCGCCTTTAAGGTTACAGCAAGAGCGCTCACGAAGACGTTCAAGTGTGCC 958
Db 999 TAACAAGAATCGCCTTTAAGGTTACAGCAAGAGCGCTCACGAAGACGTTCAAGTGTGCC 1058
Qy 959 CGTGGGACGTTTTTGAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGC 1018
Db 1059 CGTGGGACGTTTTTGAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGC 1118
Qy 1019 TCAAGCATCTTAGGGCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCC 1078
Db 1119 TCAAGCATCTTAGGGCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCC 1178
Qy 1079 CAGCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCCGCTTCACTGGCGGGA 1138
Db 1179 CAGCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCCGCTTCACTGGCGGGA 1238
Qy 1139 AGACCTCTGCAAGGGGGAAGGCAAGGCGCCCTTCAAGGCCCAACAGCGCGGTTCCGCGGC 1198
Db 1239 AGACCTCTGCAAGGGGGAAGGCAAGGCGCCCTTCAAGGCCCAACAGCGCGGTTCCGCGGC 1298
Qy 1199 CCATTGTGTAGCTTCCCGGCCACGGGCACCTCTGTCTTCCACACCTGTCTGTGTGGG 1258
Db 1299 CCATTGTGTAGCTTCCCGGCCACGGGCACCTCTGTCTTCCACACCTGTCTGTGTGGG 1358
Qy 1259 CTGTCCGGGAAGACCTAACCCTGTGGGCACTCAAGGTGTGCCCAAGCCGGCCCTGTGCTC 1318
Db 1359 CTGTCCGGGAAGACCTAACCCTGTGGGCACTCAAGGTGTGCCCAAGCCGGCCCTGTGCTC 1418

Qy 1319 AGGAGACCTCAGGGTTCTGAGATTCCTGCACTGGAACCTCATGCCCCGCTCCCA 1378
Db 1419 AGGAGACCTCAGGGTTCTGAGATTCCTGCACTGGAACCTCATGCCCCGCTCCCA 1478
Qy 1379 CGACCTGACGTAGAGGGCCCTGTGTTCCGCCATTAATTTGACAGAGCTGCTGG 1438
Db 1479 CGACCTGACGTAGAGGGCCCTGTGTTCCGCCATTAATTTGACAGAGCTGCTGG 1538
Qy 1439 TCCGTGCATATATCCAGAGGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGG 1498
Db 1539 TCCGTGCATATATCCAGAGGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGG 1598
Qy 1499 AGCGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCA 1558
Db 1599 AGCGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCCCTTCA 1658
Qy 1559 GAGCTAAGGACGAACAGCGGTGTCTCCCACTCAGGGCCTTGCTGTGGGCTTCACT 1618
Db 1659 GAGCTAAGGACGAACAGCGGTGTCTCCCACTCAGGGCCTTGCTGTGGGCTTCACT 1718
Qy 1619 TGGAAATTTCTCAGTTCCCTCCAGGCTTCTAGAACATCTGGGCCAGGGCTCATGGCTGG 1678
Db 1719 TGGAAATTTCTCAGTTCCCTCCAGGCTTCTAGAACATCTGGGCCAGGGCTCATGGCTGG 1778
Qy 1679 ATAATTTCCCTAGGCTTAAACAACCCAGCAAGCTTCGCGTCTCGTTTAATTTTGGTTA 1738
Db 1779 ATAATTTCCCTAGGCTTAAACAACCCAGCAAGCTTCGCGTCTCGTTTAATTTTGGTTA 1838
Qy 1739 AACTTATGAAATGTATTAAGAAAGAGTGCAGCTTGAAGAGATTCAAGATGGAACACA 1798
Db 1839 AACTTATGAAATGTATTAAGAAAGAGTGCAGCTTGAAGAGATTCAAGATGGAACACA 1898
Qy 1799 CCAAGACCCAGATCAAAAGCCCAACCATGCCAGCCCCCTCCAGCAGCCCCCAGCAG 1858
Db 1899 CCAAGACCCAGATCAAAAGCCCAACCATGCCAGCCCCCTCCAGCAGCCCCCAGCAG 1958
Qy 1859 ACCATGTTCTGAATTCGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGAA 1918
Db 1959 ACCATGTTCTGAATTCGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGAA 2018
Qy 1919 GGATTAACCTACCTTACGTTTTGAATAAATGTTCTCTGTTGAATG 1964
Db 2019 GGATTAACCACTTCACTGTTTGAATAAATGTTCTCTGTTGAATG 2064

RESULT 7
AAA15001
ID AAA15001 standard; cDNA; 1993 BP.
XX AC
XX AAA15001;
XX DT 21-AUG-2000 (first entry)
XX DE cDNA encoding a human proliferation and apoptosis related protein.
XX KW Human; proliferation and apoptosis related protein; PROAP; psoriasis;
XX KW cell proliferative disorder; immunological disorder; hepatitis;
XX KW reproductive disorder; arteriosclerosis; cirrhosis; lymphoma;
XX KW cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia;
XX KW asthma; diabetes mellitus; osteoarthritis; endometriosis;
XX KW uterine fibroid; menstrual cycle; ss.
XX OS Homo sapiens.
XX FH
XX Key Location/Qualifiers
FH CDS 42..1691
FT /*tag= a
FT /product= "proliferation and apoptosis related protein"
XX PN MO200023589-A2.
XX PD 27-APR-2000.

XX 19-OCT-1999; 99WO-US024511.
XX
XX 20-OCT-1998; 98US-0172216P.
PR 04-FEB-1999; 99US-0118559P.
PR 11-FEB-1999; 99US-0172229P.
PR 22-APR-1999; 99US-0154336P.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lai P;
PI Azimzal Y, Baughn MR, Yang J, Shih LL;
XX
DR WPI: 2000-339688/29.
DR P-PSDB; AAY84901.
XX
PT New human proliferation and apoptosis related protein polypeptides used
PT for diagnosis, treatment and prevention of cell proliferative,
PT immunological and reproductive disorders.
XX
XX
XX Claim 9; Page 117-118; 128pp; English.
XX
CC The present sequence encodes a human proliferation and apoptosis related
CC protein (PROAP). The polypeptides and polynucleotides can be used for the
CC diagnosis, treatment and prevention of cell proliferative, immunological
CC and reproductive disorders. Disorders associated with decreased
CC expression or activity of include arteriosclerosis, cirrhosis, hepatitis,
CC psoriasis, melanoma, lymphoma and cancers of the breast, brain and
CC prostate, acquired immune deficiency syndrome (AIDS), allergies, anaemia,
CC asthma, diabetes mellitus, osteoarthritis, endometriosis, uterine
CC fibroids and disruptions of the menstrual cycle. Antibodies against PROAP
CC can be use in diagnosis of disorders characterized by PROAP e.g. in ELISA
CC (enzyme linked immunosorbent assays) and the polynucleotides may be used
CC to detect and quantify gene expression in biopsied tissues. These
CC techniques can also be used to monitor regulation of PROAP levels during
CC therapeutic intervention
XX
SQ Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Query Match 98.3%; Score 1929.8; DB 3; Length 1993;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 ATGACGTGTTAGAGGTGCGGGCAGTTGGTGGGCACAAGAGCGAGAGCATCTATG 60
DB 42 ATGACGTGTTAGAGGTGCGGGCAGTTGGTGGGCACAAGAGCGAGAGCATCTATG 101
QY 61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGCCTTAAGCCTTTTGA 120
DB 102 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGCCTTAAGCCTTTTGA 161
QY 121 AGCTACACAACAACGTCGATCACTTTGGGATTGTACATGAGACGAGCTGCCTCTCTG 180
DB 162 AGCTACACAACAACGTCGATCACTTTGGGATTGTACATGAGACGAGCTGCCTCTCTG 221
QY 181 ACTGCCGGGAGAGCGGAAGCAATTCCGGCGGAGATCAAGCCGAAAGAGCAAGTGGTGAT 240
DB 222 ACTGCCGGGAGAGCGGAAGCAATTCCGGCGGAGATCAAGCCGAAAGAGCAAGTGGTGAT 281
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATAGATGAGCGTACAAG 300
DB 282 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGAGAAAGCTCATAGATGAGCGTACAAG 341
QY 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 360
DB 342 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 401
QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCACTTGAG 420
DB 402 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCACTTGAG 461
QY 421 CACATCCAGCGCATGACCCGGAGCTAAGCGGAGCATTTAAGGAAGCATATATTCTTCAG 480

DB 462 CACATCCAGCGCATGACCCGGAGCGTAAGCGGAGCATTTAAGGAAGCATATATTCTTCAG 521
QY 481 GATCGATACGGAAACCAAGACCGGGAACTACTCTCCATCATCTCTCTGGCATATGAGAGTAT 540
DB 522 GATCGATACGGAAACCAAGACCGGGAACTACTCTCCATCATCTCTCTGGCATATGAGAGTAT 581
QY 541 AACCCGGAGGTGGGCTACTGACAGGGAACCTGAGCCACATGCGCGCTTGTCTCTCTAT 600
DB 582 AACCCGGAGGTGGGCTACTGACAGGGAACCTGAGCCACATGCGCGCTTGTCTCTCTAT 641
QY 601 CTTCTGAGGAGATGCACTTCTGGGACCTGTGCACTGTGCTGCGCAGTGAAGGCACTCC 660
DB 642 CTTCTGAGGAGATGCACTTCTGGGACCTGTGCACTGTGCTGCGCAGTGAAGGCACTCC 701
QY 661 CTGCAAGGATTTCACAGCCCAATGCGGGACCGTCCAGGGGCTCCAGACCAAGAGAG 720
DB 702 CTGCAAGGATTTCACAGCCCAATGCGGGACCGTCCAGGGGCTCCAGACCAAGAGAG 761
QY 721 CATGTGTAGCCACGTCAACAACCCAAAGACCATGGGGCATCAGGACAGAAAGATCTATGT 780
DB 762 CATGTGTAGCCACGTCAACAACCCAAAGACCATGGGGCATCAGGACAGAAAGATCTATGT 821
QY 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
DB 822 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 881
QY 841 CTCACCTGCGCCTGTGGAGCTGTATCTGTGTAAGGCGAAACAGGCGTTGATGCCGATA 900
DB 882 CTCACCTGCGCCTGTGGAGCTGTATCTGTGTAAGGCGAAACAGGCGTTGATGCCGATA 941
QY 901 ACAAGAATCGCCTTAAGTTACAGAGAAGCGCTCAAGAGACGTCCAGGTGGCCCG 960
DB 942 ACAAGAATCGCCTTAAGTTACAGAGAAGCGCTCAAGAGACGTCCAGGTGGCCCG 1001
QY 961 TGGGCAAGTTTTCGCAACCGGTTCTGTGATCCTGGGCCAGGAGTGAAGACACTGTGCTC 1020
DB 1002 TGGGCAAGTTTTCGCAACCGGTTCTGTGATCCTGGGCCAGGAGTGAAGACACTGTGCTC 1061
QY 1021 AAGCATCTTAGGGCCTCTATGAAGAACTAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
DB 1062 AAGCATCTTAGGGCCTCTATGAAGAACTAACAAGAAAGCAGGGGACCTGCCACCCCA 1121
QY 1081 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCCGGCTTCACTGTGGGAG 1140
DB 1122 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCCGGCTTCACTGTGGGAG 1181
QY 1141 ACCCTTGCAAGGGGAGCAGGCAAGGCCCTCCAGGCCCAAGCAGCCCGTTCCCGGGCCC 1200
DB 1182 ACCCTTGCAAGGGGAGCAGGCAAGGCCCTCCAGGCCCAAGCAGCCCGTTCCCGGGCCC 1241
QY 1201 ATTTGTCAGCTTCCCGCCACGGGCACTCTCTTCCACACACCTGTCTGTGGGCT 1260
DB 1242 ATTTGTCAGCTTCCCGCCACGGGCACTCTCTTCCACACACCTGTCTGTGGGCT 1301
QY 1261 GTCCGGGAAGACACCTTACCTGTGGGCACTCAGGGTGTGCCAGCCCGCCCTGGCTAG 1320
DB 1302 GTCCGGGAAGACACCTTACCTGTGGGCACTCAGGGTGTGCCAGCCCGCCCTGGCTAG 1361
QY 1321 GAGAGACCTCAGGGTTCCTGAGATTCTGTGAGTGAACCTCATGCCCGCTCCCAAG 1380
DB 1362 GAGAGACCTCAGGGTTCCTGAGATTCTGTGAGTGAACCTCATGCCCGCTCCCAAG 1421
QY 1381 GACCTGACGTAGAGGGCCTTGGTTCCGCCATTATGATTTACAGACAGAGCTGTGGGTC 1440
DB 1422 GACCTGACGTAGAGGGCCTTGGTTCCGCCATTATGATTTACAGACAGAGCTGTGGGTC 1481
QY 1441 CGTGCCATATCCAGAGAGACCAAGTGGCCCTGCTGGCAGGCTGAACAACCTTGGGAG 1500
DB 1482 CGTGCCATATCCAGAGAGACCAAGTGGCCCTGCTGGCAGGCTGAACAACCTTGGGAG 1541
QY 1501 CGGGTGAAGATCGGCTTTCGCTGCAACCCAGACCTGATTTCCGACCAAGGCAACCCCTTCAG 1560
DB 1542 CGGGTGAAGATCGGCTTTCGCTGCAACCCAGACCTGATTTCCGACCAAGGCAACCCCTTCAG 1601

Qy 1561 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCACTTG 1620
Db 1602 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCACTTG 1661
Qy 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT 1680
Db 1662 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT 1721
Qy 1681 AATTTCCTAGGCTTAACAACCCAGCAAGCTTCGCTCTGTTTATTGTTAA 1740
Db 1722 AATTTCCTAGGCTTAACAACCCAGCAAG-TTGCACTCTGTTTATTGTTAA 1780
Qy 1741 CTTATGAAAATGTATTAGAAAGAGTGACGCTCGAGAGAGATTCAAGATGAACACACC 1800
Db 1781 CTTATGAAAATGTATTAGAAAGAGTGACGCTCGAGAGAGATTCAAGATGAACACACC 1840
Qy 1801 AGACCCGAGATCAACAAGCCAAAGCCATGCCAGCCCTCCCAACACCCCGCCAGCCACGAC 1860
Db 1841 AGACCCGAGATCAACAAGCCAAAGCCATGCCAGCCCTCCCAACACCCCGCCAGCCACGAC 1900
Qy 1861 CATCGTTCTGAATTCTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
Db 1901 CATCGTTCTGAATTCTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1960
Qy 1921 ATAACACTCTTCAACGTTTGAATAATGTTTC 1953
Db 1961 ATAACACTCTTCAATGTTTGAATAATGGGTC 1993

RESULT 8

ADT88069
ID ADT88069 standard; cDNA; 1993 BP.

AC ADT88069;

DT 30-DEC-2004 (first entry)

DE Human proliferation and apoptosis related protein (PROAP)-12 cDNA.

XX PROAP; proliferation and apoptosis related protein;
KW cell proliferative disorder; cancer; atherosclerosis;
KW immunological disorder; AIDS; acquired immunodeficiency syndrome;
KW allergy; reproductive disorder; infertility; gene therapy; cytostatic;
KW antiarteriosclerotic; immunosuppressive; anti-HIV; antiallergic;
KW antiinfertility; gynaecological; human; gene; ss.

XX Homo sapiens.

OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 42..1691
FT /*tag= a
FT /product= "Proliferation and apoptosis related protein
FT (PROAP)-12"

PN US2004203106-A1.

XX 14-OCT-2004.

XX 05-MAY-2004; 2004US-00839882.

XX 19-JAN-1999; 99US-0172216P.
XX 04-FEB-1999; 99US-0118559P.
XX 11-FEB-1999; 99US-0172229P.
XX 22-APR-1999; 99US-0154336P.
XX 19-OCT-1999; 99WO-US024511.
XX 11-APR-2001; 2001US-00807452.

XX (INCY-) INCYTE CORP.

XX Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lai P,
PI Azimzai Y, Baughn MR, Yang J, Shih LL,
XX

DR WPI; 2004-728011/71.
DR P-PSDB; ADT88050.

XX New human proliferation and apoptosis related proteins and
PT polynucleotides for diagnosing, preventing or treating disorders
PT associated with aberrant protein expression, e.g. cancer, AIDS,
PT atherosclerosis or infertility.

XX Claim 5; SEQ ID NO 31; 85pp; English.

XX The present invention relates to the human proliferation and apoptosis
CC related protein (PROAP) and its encoding nucleic acid. The invention is
CC useful for diagnosing, preventing or treating disorders associated with
CC altered expression or activity of human PROAP, such as cell proliferative
CC (e.g. cancer or atherosclerosis), immunological (e.g. acquired
CC immunodeficiency syndrome (AIDS) or allergies) or reproductive (e.g.
CC infertility) disorders. The invention is also be used in screening for
CC drugs that may be used for treating or preventing the disorders mentioned
CC above and in gene therapy. The present sequence is the human PROAP-12
CC protein encoding cDNA.

XX Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Qy Query Match 98.3%; Score 1929.8; DB 13; Length 1993;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 1 ATGAGCTGTGTAGAGTGGCGGCAAGTTGTTGGGCAAGAGCGAGAGACATCATATG 60
Db 42 ATGAGCTGTGTAGAGTGGCGGCAAGTTGTTGGGCAAGAGCGAGAGACATCATATG 101

Qy 61 AAATACGAAAAGGACACCCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTCA 120
Db 102 AAATACGAAAAGGACACCCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTCA 161

Qy 121 AGCTACACAACAACGTCATCTTTGGGATTGTACATGAGACGAGCTGCTCTCTG 180
Db 162 AGCTACACAACAACGTCATCTTTGGGATTGTACATGAGACGAGCTGCTCTCTG 221

Qy 181 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240
Db 222 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 281

Qy 241 ATGCTGGAGAGTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGGTACAAG 300
Db 282 ATGCTGGAGAGTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGGTACAAG 341

Qy 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAAACATTAAGAAATG 360
Db 342 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAAACATTAAGAAATG 401

Qy 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGGCAAGAGTCACTGAG 420
Db 402 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGGCAAGAGTCACTGAG 461

Qy 421 CACATCCAGCGCATCGACGGGACGTAAGCGGACATTAAGAGCATATATTCTTCAGG 480
Db 462 CACATCCAGCGCATCGACGGGACGTAAGCGGACATTAAGAGCATATATTCTTCAGG 521

Qy 481 GATGATACGGAACCAAGCAGCGGAACTACTCCATCTCTGGCATTAAGAGATAT 540
Db 522 GATGATACGGAACCAAGCAGCGGAACTACTCCATCTCTGGCATTAAGAGATAT 581

Qy 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATCGCCGCTTGTCTCTCTAT 600
Db 582 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCATCGCCGCTTGTCTCTCTAT 641

Qy 601 CTTCTGAGGAGGATGATTTCTGGGCACTGTGTGAGCTGTGGCCAGTGAAGGCACTCC 660
Db 642 CTTCTGAGGAGGATGATTTCTGGGCACTGTGTGAGCTGTGGCCAGTGAAGGCACTCC 701

Qy 661 CTGCAAGGATTTACAGCCCAATGCGGGAGCCGTCCAGGGGCTCCAAGACCAACAGAG 720

Db 702 CTGCAAGGATTTCACAGCCCCAAATGCGGGGACCGTCCAGGGGCTCCAAGACCAACAGAG 761
Qy 721 CATGTGTAGCCACGTCACAACCCCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 780
Db 762 CATGTGTAGCCACGTCACAACCCCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 821
Qy 781 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
Db 822 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 881
Qy 841 CTCACCCCTGGCCTGTGGGACGTGTATCTGTGTAAGAGCGAACAGGCGTTGATGCCGATA 900
Db 882 CTCACCCCTGGCCTGTGGGACGTGTATCTGTGTAAGAGCGAACAGGCGTTGATGCCGATA 941
Qy 901 ACAAGAATCGCCTTTAAGGTTACAGCAGAAAGCGCCTCACGAAGACGTCCAGGTGTGGCCG 960
Db 942 ACAAGAATCGCCTTTAAGGTTACAGCAGAAAGCGCCTCACGAAGACGTCCAGGTGTGGCCG 1001
Qy 961 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1020
Db 1002 TGGGCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1061
Qy 1021 AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1062 AAGCATCTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGACCTGCCACCCCA 1121
Qy 1081 GCCAAACCCGACCAAGGGTCTGCGGATCCAGGCTGTGCGGCTTCACTGGCGGGAAG 1140
Db 1122 GCCAAACCCGACCAAGGGTCTGCGGATCCAGGCTGTGCGGCTTCACTGGCGGGAAG 1181
Qy 1141 ACCCTCTGCAAGGGGGACAGGAGGCCCCCTCCAGGCCCAACGCCCCGTTCCCGGCGCC 1200
Db 1182 ACCCTCTGCAAGGGGGACAGGAGGCCCCCTCCAGGCCCAACGCCCCGTTCCCGGCGCC 1241
Qy 1201 ATTTGGTCAAGCTTCCCCCGCACGGGACCTGTTCTTCCACACCTGTCTGTGGGGCT 1260
Db 1242 ATTTGGTCAAGCTTCCCCCGCACGGGACCTGTTCTTCCACACCTGTCTGTGGGGCT 1301
Qy 1261 GTCCGGGAAGACCTTACCCTGTGGGCACTAGGGTGTGCCAGGCCCGGCTGTGGCTCAG 1320
Db 1302 GTCCGGGAAGACCTTACCCTGTGGGCACTAGGGTGTGCCAGGCCCGGCTGTGGCTCAG 1361
Qy 1321 GGAGGACCTCAGGGTCTCTGAGATTCTGCACTGGAATCCCATGCCCCGCTCCCAAG 1380
Db 1362 GGAGGACCTCAGGGTCTCTGAGATTCTGCACTGGAATCCCATGCCCCGCTCCCAAG 1421
Qy 1381 GACCTGACGTAGAGGGCCCTTGGTCCGCCATTATGATTTCAACAAGAGCTGTGGGTC 1440
Db 1422 GACCTGACGTAGAGGGCCCTTGGTCCGCCATTATGATTTCAACAAGAGCTGTGGGTC 1481
Qy 1441 CGTGCCATATCCCAAGAGGACCAAGCTGGGCCCCCTGCTGGCAAGCTGAACACCCCTGCCAG 1500
Db 1482 CGTGCCATATCCCAAGAGGACCAAGCTGGGCCCCCTGCTGGCAAGCTGAACACCCCTGCCAG 1541
Qy 1501 CGGTGAGATCGGCTTTGCTGTCACCCAGCACTGATTCGACCAAGGACCCCTTCAGA 1560
Db 1542 CGGTGAGATCGGCTTTGCTGTCACCCAGCACTGATTCGACCAAGGACCCCTTCAGA 1601
Qy 1561 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCTTGCCTTGCAGGCTCCACTTG 1620
Db 1602 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCTTGCCTTGCAGGCTCCACTTG 1661
Qy 1621 GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAAGCATCTGGGCAAGGCTCATGGCTGAT 1680
Db 1662 GAAAGTTCTCAGTTCCCTCCAGGCTTCTAGAAGCATCTGGGCAAGGCTCATGGCTGAT 1721
Qy 1681 AATTTCCTTAGGCTTAACAACCCCAAGCAAGCTTGCCTCTCTGTTTATTTTGGTTAAA 1740
Db 1722 AATTTCCTTAGGCTTAACAACCCCAAGCAAG-TTCGCATCTCTGTTTATTTTGGTTAAA 1780
Qy 1741 CTTATGAAAATGTATTAAGAAAGATGAGCTCGAGAGAGATTCAGAGATGGAACACACC 1800
Db 1781 CTTATGAAAATGTATTAAGAAAGATGAGCTCGAGAGAGATTCAGAGATGGAACACACC 1840

Qy 1801 AGACCCGAGATCACAAAGCCAACCATGCCAGGCCCTCCAGACACCCCCAGCCCCAGAC 1860
Db 1841 AGACCCGAGATCACAAAGCCAACCATGCCAGGCCCTCCAGACACCCCCAGCCCCAGAC 1900
Qy 1861 CATCGTTCTGAATTCTGACGACACCCGTGAGCCCTGCTTTGTACTTTAACTCATGGAAG 1920
Db 1901 CATCGTTCTGAATTCTGACGACACCCGTGAGCCCTGCTTTGTACTTTAACTCATGGAAG 1960
Qy 1921 ATAACCTCTTACGTTTGAATAATGTTTC 1953
Db 1961 ATAACCACTTCATGTTTGAATAATGGGTC 1993

RESULT 9
ADC37380
ID ADC37380 standard; DNA; 2304 BP.
XX
AC ADC37380;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 213.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN MO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
DR P-PSDB; ADC37381.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 4; SEQ ID NO 213; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischaemic disorders.
XX
SQ Sequence 2304 BP; 569 A; 610 C; 636 G; 489 T; 0 U; 0 Other;
Query Match 96.8%; Score 1902; DB 10; Length 2304;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1937; Conservative 0; Mismatches 5; Indels 22; Gaps 1;

Qy 1 ATGACGTGTAGAGGTGCGGGGCACTGTGGGCACAGCGAGAGCATCATATG 60
Db 362 ATGACGTGTAGAGGTGCGGGGCACTGTGGGCACAGCGAGAGCATCATATG 421

QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTCGA 120
DB 422 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTCGA 481
QY 121 AGCTACAAACAACAGCTGCATCTTTGGGGATTGTACATGAGACGAGCTGCCTCTCTG 180
DB 482 AGCTACAAACAACAGCTGCATCTTTGGGGATTGTACATGAGACGAGCTGCCTCTCTG 541
QY 181 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 240
DB 542 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 601
QY 241 ATGCTGGAGACTGGAGAAATACAAAGCAGCAAAAGCTCATAGATCGAGCGTACAAG 300
DB 602 ATGCTGGAGACTGGAGAAATACAAAGCAGCAAAAGCTCATAGATCGAGCGTACAAG 661
QY 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACATTGAGGAATG 360
DB 662 GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACATTGAGGAATG 721
QY 361 AAGTTGAAAAACCCCGGAAGATACCATCATGAAGGAAAGGCAAGAGTCACTGAG 420
DB 722 AAGTTGAAAAACCCCGGAAGATACCATCATGAAGGAAAGGCAAGAGTCACTGAG 781
QY 421 CACATCCAGCGCATCGACCCGGAGCTAAGCGGACATTAAGGAAGCATATATTCTCAAG 480
DB 782 CACATCCAGCGCATCGACCCGGAGCTAAGCGGACATTAAGGAAGCATATATTCTCAAG 841
QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCCTCTGGCATATGAGAGTAT 540
DB 842 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCCTCTGGCATATGAGAGTAT 901
QY 541 AACC CGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 600
DB 902 AACC CGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 961
QY 601 CTTCTGAGAGAGATGCACTTCTGGGCACTGGTGACGTGCTGGCCAGTGAGAGCACTCC 660
DB 962 CTTCTGAGAGAGATGCACTTCTGGGCACTGGTGACGTGCTGGCCAGTGAGAGCACTCC 1021
QY 661 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 720
DB 1022 CTGCAAGGATTTCAAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 1081
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGAGGCACTCAGACAAGAAAGATCTATGT 780
DB 1082 CATGTGTAGCCACGTCAACAACCAAGACCATGAGGCACTCAGACAAGAAAGATCTATGT 1141
QY 781 GGGCAGTGTCCCCCGTTAGGCTCCTCATCCGGATATTGATTGACGGGATCTCTCGGG 840
DB 1142 GGGCAGTGTCCCCCGTTAGGCTCCTCATCCGGATATTGATTGACGGGATCTCTCGGG 1201
QY 841 CTCACCCCTGCGCCTGTGGAGCGTATCTGTAGAAAGCGAACAAGCGTTGATGCCGATA 900
DB 1202 CTCACCCCTGCGCCTGTGGAGCGTATCTGTAGAAAGCGAACAAGCGTTGATGCCGATA 1261
QY 901 ACAAGATCGCCTTTAAGTTCAAGAGAAGCGCCTCAAGAAAGCTCCAGGTGTGCCCG 960
DB 1262 ACAAGATCGCCTTTAAGTTCAAGAGAAGCGCCTCAAGAAAGCTCCAGGTGTGCCCG 1321
QY 961 TGGGCACTTTTTCACACCGGTTCTGTATACCTGGCCAGGGATGAGAGCACTGTGCTC 1020
DB 1322 TGGGCACTTTTTCACACCGGTTCTGTATACCTGGCCAGGGATGAGAGCACTGTGCTC 1381
QY 1021 AAGCATCTTAGGGCTCTATGAAGAAATAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
DB 1382 AAGCATCTTAGGGCTCTATGAAGAAATAACAAGAAAGCAGGGGACCTGCCACCCCA 1441
QY 1081 GCCAAACCCGAGCAAGGCTCTGCGCATCCAGGCTGTCCGGCTTACGTCGCGGGAAG 1140
DB 1442 GCCAAACCCGAGCAAGGCTCTGCGCATCCAGGCTGTCCGGCTTACGTCGCGGGAAG 1501

QY 1141 ACCCTGCAAGGGGACAGGACAGGCCCTCTGAGGCCCAACAGCCGGTTCCGCGGCC 1200
DB 1502 ACCCTGCAAGGGGACAGGACAGGCCCTCTCA-----GGCCC 1539
QY 1201 ATTTGTCACTTCCCGCCACGGGCACTCTGTTCTTCCACACCCTGTCTGTGGGCT 1260
DB 1540 ATTTGTCACTTCCCGCCACGGGCACTCTGTTCTTCCACACCCTGTCTGTGGGCT 1599
QY 1261 GTCCGGGAAGACACCTACCCTGTGGGCACTCAGGGTGTCCCAAGCCCGGCTGTGCTAG 1320
DB 1600 GTCCGGGAAGACACCTACCCTGTGGGCACTCAGGGTGTCCCAAGCCCGGCTGTGCTAG 1659
QY 1321 GGAGCACTCAGGGTCTCTGAGATCTCTGAGTGAATCCATGCCCCGCTGCCAACG 1380
DB 1660 GGAGCACTCAGGGTCTCTGAGATCTCTGAGTGAATCCATGCCCCGCTGCCAACG 1719
QY 1381 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 1440
DB 1720 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTCAAGACAGAGCTGTGGGTC 1779
QY 1441 CGTGCCATATCCAGAGACCAAGCTGGCCCCCTGCTGGCAAGGCTGAACACCTGCGGAG 1500
DB 1780 CGTGCCATATCCAGAGACCAAGCTGGCCCCCTGCTGGCAAGGCTGAACACCTGCGGAG 1839
QY 1501 CGGTTGACATCGGCTTTCGCTGCAACCCAGCACTGATTTCCGACAGGGCACCCCTTCAGA 1560
DB 1840 CGGTTGACATCGGCTTTCGCTGCAACCCAGCACTGATTTCCGACAGGGCACCCCTTCAGA 1899
QY 1561 GCTAGGACGAACAGCCGTGTGCTCCACCTCAGGGCTTGCTGCGGCTTCACTTG 1620
DB 1900 GCTAGGACGAACAGCCGTGTGCTCCACCTCAGGGCTTGCTGCGGCTTCACTTG 1959
QY 1621 GAAAGTTCTCACTTCCCTCCAGGCTTTAGAAAGCATGTGGCCAGGGCTCATGGCTGAT 1680
DB 1960 GAAAGTTCTCACTTCCCTCCAGGCTTTAGAAAGCATGTGGCCAGGGCTCATGGCTGAT 2019
QY 1681 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCTCTGTTTATTTGGTTAAA 1740
DB 2020 AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCTCTGTTTATTTGGTTAAA 2079
QY 1741 CTTATGAATAATGTATTAAGAAAGTGACGCTCGAGAGAGATTCAAGATGGAACACACC 1800
DB 2080 CTTATGAATAATGTATTAAGAAAGTGACGCTCGAGAGAGATTCAAGATGGAACACACC 2139
QY 1801 AGACCCAGATCACAAAGCCAACCATGCCAGCCCCCTCCAGCACCCCCCAGCCACGAC 1860
DB 2140 AGACCCAGATCACAAAGCCAACCATGCCAGCCCCCTCCAGCACCCCCCAGCCACGAC 2199
QY 1861 CATGTTCTGAATTCAGACACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAGG 1920
DB 2200 CATGTTCTGAATTCAGACACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAGG 2259
QY 1921 ATAACCTACTCAGCTTTTGAATAAATGTTTCTGTGAATG 1964
DB 2260 ATAACCTACTCAGCTTTTGAATAAATGTTTCTGTGAATG 2303

RESULT 10
ADC37378 standard; DNA; 2304 BP.
XX
AC ADC37378;
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 211.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX

OS	Homo sapiens.
XX	
PN	WO2003048202-A2.
XX	
PD	12-JUN-2003.
XX	
PF	03-DEC-2002; 2002WO-JP012644.
XX	
PR	03-DEC-2001; 2001JP-00368692.
PR	05-DEC-2001; 2001US-0335829P.
PR	03-OCT-2002; 2002JP-00291302.
PR	04-OCT-2002; 2002US-0415769P.
XX	
PA	(ASAH) ASAMI KASEI KK.
XX	
PI	Matsuda A, Muramatsu S;
XX	
DR	WPI; 2003-505282/47.
DR	P-PSDB; ADC37379.
XX	
PT	New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT	useful for treating inflammation, autoimmune diseases, cancers,
PT	infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX	ischemic disorders.
PS	Claim 4; SEQ ID NO 211; 938pp; English.

The present invention relates to novel proteins and their coding sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-kappaB). The proteins and their coding sequences are useful for treating a disease associated with NF-kappaB activation, such as inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases, or ischaemic disorders.

Sequence 2304 BP; 570 A; 609 C; 636 G; 489 T; 0 U; 0 Other;

Query Match	96.8%	Score 1900.4;	DB 10;	Length 2304;
Best Local Similarity	98.6%	Pred. No. 0;		
Matches 1936; Conservative	0;	Mismatches	6;	Indels 22; Gaps 1;

QY	1	ATGACGCTGTAGAGGTCGCGGGCAGTTGGTGGGCACAAAGCCAGAGGACATCATTATG	60
Db	362	ATGACGCTGTAGAGGTCGCGGGCAGTTGGTGGGCACAAAGCCAGAGGACATCATTATG	421
QY	61	AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA	120
Db	422	AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA	481
QY	121	AGCTACAACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCCTCTG	180
Db	482	AGCTACAACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCCTCTG	541
QY	181	ACTGCGCGGGAGGCGAAGCBAATTCCGCGGGAGATCAGCCGAAGAAGCAAGTGGTGAT	240
Db	542	ACTGCGCGGGAGGCGAAGCBAATTCCGCGGGAGATCAGCCGAAGAAGCAAGTGGTGAT	601
QY	241	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAAGCTCATAGATCGAGCGTACAAG	300
Db	602	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAAGCTCATAGATCGAGCGTACAAG	661
QY	301	GGAAATGCCCATGAAATCCCGGGGCCCGATGTGTCAGTCCCTCTGAACATTGAGGAAATG	360
Db	662	GGAAATGCCCATGAAATCCCGGGGCCCGATGTGTCAGTCCCTCTGAACACTGAGGAAATG	721
QY	361	AAAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAGAGGGCAAGAGTCATCTGAG	420
Db	722	AAAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAGAGGGCAAGAGTCATCTGAG	781
QY	421	CACATCCAGCGCATGACCCGGGACGTAAAGCGGGACATTAAAGGAAGTATATTCTTCAGG	480
Db	782	CACATCCAGCGCATGACCCGGGACGTAAAGCGGGACATTAAAGGAAGTATATTCTTCAGG	841
QY	481	GATGATACGGAACCAAGCAGCGGGAATTAATCCACATCCTCTGCGCATATGAGAGTAT	540

Db	842	GATCGATACGGAACCAAGCAGCGGAACCTACTCCACATCTCTGGCAATAGAGATAT	901
QY	541	AACCCGAGGTGGGCTACTGCAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT	600
Db	902	AACCCGAGGTGGGCTACTGCAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT	961
QY	601	CTTCTGAGAGAGATGCATTCTGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTCC	660
Db	962	CTTCTGAGAGAGATGCATTCTGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTCC	1021
QY	661	CTGCAAGGATTTCA CAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGCAACAGAG	720
Db	1022	CTGCAAGGATTTCA CAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGCAACAGAG	1081
QY	721	CATGTGTAGCCACGTCA CAAACCCAAAGACCA TGGGGCATCAGACAAGAAAGATCTATGT	780
Db	1082	CATGTGTAGCCACGTCA CAAACCCAAAGACCA TGGGGCATCAGACAAGAAAGATCTATGT	1141
QY	781	GGGCAGTGTTCCCGCTTAGGCTGCCTCATCCGATATTGATTGACGGGATCTCTCGGG	840
Db	1142	GGGCAGTGTTCCCGCTTAGGCTGCCTCATCCGATATTGATTGACGGGATCTCTCGGG	1201
QY	841	CTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGCA CAGGCGTTGATGCCGATA	900
Db	1202	CTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGCA CAGGCGTTGATGCCGATA	1261
QY	901	ACAAGAATCGCCTTTAAGGTT CAGCAGAAAGCGCTCACGAAGACGTCCAGGTGGCCCCG	960
Db	1262	ACAAGAATCGCCTTTAAGGTT CAGCAGAAAGCGCTCACGAAGACGTCCAGGTGGCCCCG	1321
QY	961	TGGGACGTTTTTGC CAAACCGGTTCTGTGATCTGGGCCAGGGATGAGGAC TGTGCTC	1020
Db	1322	TGGGACGTTTTTGC CAAACCGGTTCTGTGATCTGGGCCAGGGATGAGGAC TGTGCTC	1381
QY	1021	AAGCATCTTAGGCGCTCTATGAAGAACTTAA CAAAGAACAGGGGACTTGCA CCCCCA	1080
Db	1382	AAGCATCTTAGGCGCTCTATGAAGAACTTAA CAAAGAACAGGGGACTTGCA CCCCCA	1441
QY	1081	GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCA CGTGGCGGGAAG	1140
Db	1442	GCCAAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCA CGTGGCGGGAAG	1501
QY	1141	ACCCTCTGCAAGGGGGA CAGGCAAGGCCCTTCCA GGGCCCA CAGGCCCGGTTCCCCGGCCC	1200
Db	1502	ACCCTCTGCAAGGGGGA CAGGCAAGGCCCTTCCA GGGCCCA CAGGCCCGGTTCCCCGGCCC	1539
QY	1201	ATTGTGTACGTTCCCGCCACCGGCACCTGCTTCCACACACCTGTCTGTGGGGCT	1260
Db	1540	ATTGTGTACGTTCCCGCCACCGGCACCTGCTTCCACACACCTGTCTGTGGGGCT	1599
QY	1261	GTCCGGGAAGACCTTACCCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTTGCTCAG	1320
Db	1600	GTCCGGGAAGACCTTACCCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTTGCTCAG	1659
QY	1321	GGAGAACCTCAGGGTTCCTGAGATTCTGCA GTGGAAC TCCATGCCCCGCTCCCAACG	1380
Db	1660	GGAGAACCTCAGGGTTCCTGAGATTCTGCA GTGGAAC TCCATGCCCCGCTCCCAACG	1719
QY	1381	GACCTGACGTAGAAGGGCCCTTGTTCGGCA TTATGATTT CAGACAGAGCTGTGGGTC	1440
Db	1720	GACCTGACGTAGAAGGGCCCTTGTTCGGCA TTATGATTT CAGACAGAGCTGTGGGTC	1779
QY	1441	CGTGCCATATCCAGGAAGAC CAGCTGGCCCCCTGCTGGCAGGCTGAACACCTTCGGAG	1500
Db	1780	CGTGCCATATCCAGGAAGAC CAGCTGGCCCCCTGCTGGCAGGCTGAACACCTTCGGAG	1839
QY	1501	CGGTGAGATCGGCTTTCGCTGCACCCAGCA CTGATTCGAC CAGGCA CCCCCTTCAGA	1560
Db	1840	CGGTGAGATCGGCTTTCGCTGCACCCAGCA CTGATTCGAC CAGGCA CCCCCTTCAGA	1899
QY	1561	GCTAGGAGCA CAGCCGTGTGCTCCACCT CAGGGCCTTGCTGTGGGCTTCACTTG	1620

Db 1900 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTGTGCGGCTTCACTTG 1959
Qy 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGCCAGGGCTCATGGCTGGAT 1680
Db 1960 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGCCAGGGCTCATGGCTGGAT 2019
Qy 1681 AATTTCCCTAGGCTTAAACAACCCAGAGCTTCGCGTCCCTGTTTATTTTGGTTAAA 1740
Db 2020 AATTTCCCTAGGCTTAAACAACCCAGAGCTTCGCGATCCCTGTTTATTTTGGTTAAA 2079
Qy 1741 CTTATGAAATGTATTAGAAAGAGTCAGCTCGAGAGATTCAAGATGGAACACACC 1800
Db 2080 CTTATGAAATGTATTAGAAAGAGTCAGCTCGAGAGATTCAAGATGGAACACACC 2139
Qy 1801 AGACCCCAAGTACAAAGCCAAACCATGCCCCCTCCAGACACCCCAAGCCCAAGAC 1860
Db 2140 AGACCCCAAGTACAAAGCCAAACCATGCCCCCTCCAGACACCCCAAGCCCAAGAC 2199
Qy 1861 CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
Db 2200 CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 2259
Qy 1921 ATAACCTCTCAAGCTTTTGAATTAATGTTTCTGTTGAATG 1964
Db 2260 ATAACCTCTCAAGCTTTTGAATTAATGTTTCTGTTGAATG 2303

RESULT 11
ABZ11810
ID ABZ11810 standard; cDNA; 2084 BP.
XX AC ABZ11810;
XX DT 20-JAN-2003 (first entry)
DE Human polynucleotide SEQ ID NO 692.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
antiarthritis; gene; ss.
KW
XX
OS Homo sapiens.
XX
PN WO200270539-A2.
XX
PD 12-SEP-2002.
XX
PF 05-MAR-2002; 2002WO-US005095.
XX
PR 05-MAR-2001; 2001US-00799451.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2002-759812/82.
DR P-PSDB; ABP69593.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for creating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
PS Claim 1; SEQ ID NO 692; 1012pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burn, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from Wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Query Match 93.5%; Score 1836.4; DB 6; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

Qy 1 ATGACGTGTAGAGTCCGCGGCAAGTGTGGGCAACAAGCGAGACATCATATG 60
Db 163 ATGACGTGTAGAGTCCGCGGCAAGTGTGGGCAACAAGCGAGACATCATATG 222
Qy 61 AATATCGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTAAGCCTTTTGA 120
Db 223 AATATCGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTAAGCCTTTTGA 282
Qy 121 AGCTACAAACAACGTCATCTTGGGATTGTATAGACGAGAGCTGCTCTCTG 180
Db 283 AGCTACAAACAACGTCATCTTGGGATTGTATAGACGAGAGCTGCTCTCTG 342
Qy 181 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT 240
Db 343 ACTGCGGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGAGCAAGTGGTGAT 402
Qy 241 ATGCTGGGAGACTGGGAAATTAACAAAGCAGACAAAGCTCATAGATCGAGCTACAAG 300
Db 403 ATGCTGGGAGACTGGGAAATTAACAAAGCAGACAAAGCTCATAGATCGAGCTACAAG 462
Qy 301 GGAATGCCATGAACATCCGGGCCCCGATGTGTCTCTCTGAAACATTAAGAAATG 360
Db 463 GGAATGCCATGAACATCCGGGCCCCGATGTGTCTCTCTGAAACATTAAGAAATG 522
Qy 361 AAGTTGAAAAACCCCGAAGATACCATCATGAAGAGAAAGGCAAGGTCATCTGAG 420
Db 523 AAGTTGA----- 529
Qy 421 CACATCCAGCGCATGACCGGAGCTAAGCGGACATTAAGGAAGCATATATTCTCAGG 480
Db 530 -ACATCCAGCGCATGACCGGAGCTAAGCGGACATTAAGGAAGCATATATTCTCAGG 588
Qy 481 GATGATACGGAACCAAGCAGCGGGAATCTACTCCATCTCTGSCATATGAGAGATAT 540
Db 589 GATGATACGGAACCAAGCAGCGGGAATCTACTCCATCTCTGSCATATGAGAGATAT 648
Qy 541 AACCCTGAGGTGGGTACTTGAAGGAGACCTGAGCCCATCGCCCTTGTCTCTCTAT 600
Db 649 AACCCTGAGGTGGGTACTTGAAGGAGACCTGAGCCCATCGCCCTTGTCTCTCTAT 708
Qy 601 CTTCTGAGGAGATGATCTTCTGGGCACTGTGCTGCTGCGCAGTGAAGGCACTCC 660
Db 709 CTTCTGAGGAGATGATCTTCTGGGCACTGTGCTGCTGCGCAGTGAAGGCACTCC 768
Qy 661 CTGCAAGGATTTACAGCCCAATGGCGGAGCCGTCCAGGGGCTCAAGCAACAGAGAG 720
Db 769 CTGCAAGGATTTACAGCCCAATGGCGGAGCCGTCCAGGGGCTCAAGCAACAGAGAG 828
Qy 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGCTATCAGACAAAGAAATCTATGT 780
Db 829 CATGTGTAGCCACGTCAACAACCAAGACCATGGGCTATCAGACAAAGAAATCTATGT 888

Db 283 AGCTACACACACACGTCGATCAATTTGGGGATTGTACATGAGACGGAGCTGCTCTCTG 342
Qy 181 ACTGCGCGGAGAGCGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGACCAAGTGGTGGAT 240
Db 343 ACTGCGCGGAGAGCGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGACCAAGTGGTGGAT 402
Qy 241 ATGCTGGGAGACTGGAGAAATACAAAAGCAGCAGAAAGCTCATAGTCGAGCGTACAAG 300
Db 403 ATGCTGGGAGACTGGAGAAATACAAAAGCAGCAGAAAGCTCATAGTCGAGCGTACAAG 462
Qy 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAAATG 360
Db 463 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAAATG 522
Qy 361 AAGTTGAAAAAACCAGGAATACCATCATGAAGAGAAAGGCAAGGTCATCTGAG 420
Db 523 AAGTTGA----- 529
Qy 421 CACATCCAGCGCATCGACCCGGGACGTAAAGCGGGACATTAAGGAAGCATATATTCTTCAGG 480
Db 530 -ACATCCAGCGCATCGACCCGGGACGTAAAGCGGGACATTAAGGAAGCATATATTCTTCAGG 588
Qy 481 GATCGATACGGAACCAAGCAGCGGGAATCTACCATCTCTCTGGCATATGAGAGTAT 540
Db 589 GATCGATACGGAACCAAGCAGCGGGAATCTACCATCTCTCTGGCATATGAGAGTAT 648
Qy 541 AACCCGAGGTGGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 649 AACCCGAGGTGGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 708
Qy 601 CTTCTGAGAGGATGATGATTTCTGGGCACTGTGTGACGTCTGGCCAGTGAAGGCACTCC 660
Db 709 CTTCTGAGAGGATGATGATTTCTGGGCACTGTGTGACGTCTGGCCAGTGAAGGCACTCC 768
Qy 661 CTGCAAGGATTTTCAACAGCCCAAATGGCGGGACCGTCCAGGGGCTCCAAGACCAACAGAG 720
Db 769 CTGCAAGGATTTTCAACAGCCCAAATGGCGGGACCGTCCAGGGGCTCCAAGACCAACAGAG 828
Qy 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATGT 780
Db 829 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAAGATCTATGT 888
Qy 781 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATGATGACGGGATCTCTCGGG 840
Db 889 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATGATGACGGGATCTCTCGGG 948
Qy 841 CTCACCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGAAGAGCGGCTGATGCCGATA 900
Db 949 CTCACCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGAAGAGCGGCTGATGCCGATA 1008
Qy 901 ACAAGATCGCCTTTAAGGTTACAGACAGAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 960
Db 1009 ACAAGATCGCCTTTAAGGTTACAGACAGAGCGCCTCACGAAGACGTCCAGGTGTGCCCG 1068
Qy 961 TGGGCAAGTTTGTCAACCGGTTCTGTATACCTGGGCAAGGATGAGACACTGTGCTC 1020
Db 1069 TGGGCAAGTTTGTCAACCGGTTCTGTATACCTGGGCAAGGATGAGACACTGTGCTC 1128
Qy 1021 AAGCATCTTAGGGCCTTATGAGAAACTAACAAAGACAGGGGACCTGCCACCCCA 1080
Db 1129 AAGCATCTTAGGGCCTTATGAGAAACTAACAAAGACAGGGGACCTGCCACCCCA 1188
Qy 1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTACGTCGGGGAAG 1140
Db 1189 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTACGTCGGGGAAG 1248
Qy 1141 ACCCTCTGCAAGGGGACAGGACAGGCGCCCTCCAGGCCCAACAGCCGGTTCCCGCGGCC 1200
Db 1249 ACCCTCTGCAAGGGGACAGGACAGGCGCCCTCCAGGCCCAACAGCCGGTTCCCGCGGCC 1308
Qy 1201 AATTGTCAGCTTCCCGCCACGGGCACTCTGTTTTCACACCTGTCTGTGGGGCT 1260

Db 1309 AATTGTCAGCTTCCCGCCACGGGCACTGTGTTCTTCCACACCTGTCTGTGGGGCT 1368
Qy 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCACGCCGCTGGCTCAG 1320
Db 1369 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCACGCCGCTGGCTCAG 1428
Qy 1321 GGAGACCTCAGGGTTCCTTGAGATTCTCTGACGTGGAATCCATGCCCCGCTTCCCAAG 1380
Db 1429 GGAGACCTCAGGGTTCCTTGAGATTCTCTGACGTGGAATCCATGCCCCGCTTCCCAAG 1488
Qy 1381 GACCTGACGTAGAGGGCCCTTGGTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1440
Db 1489 GACCTGACGTAGAGGGCCCTTGGTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1548
Qy 1441 CGTCCATATCCAGAGACAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1500
Db 1549 CGTCCATATCCAGAGACAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1608
Qy 1501 CGGTCAGATCGGCTTTCGCTGCAACCCAGCATGTATTCCGACAGGCAACCCCTTCAGA 1560
Db 1609 CGGTCAGATCGGCTTTCGCTGCAACCCAGCATGTATTCCGACAGGCAACCCCTTCAGA 1668
Qy 1561 GCTAGGACGAACAGCCGTGTCTCCCACTCAGGGCTTGCCTGTGGGCTTCACTTG 1620
Db 1669 GCTAGGACGAACAGCCGTGTCTCCCACTCAGGGCTTGCCTGTGGGCTTCACTTG 1728
Qy 1621 GAAAGTTCTCAGTTCCTTCCAGGCTTCTAGAAGCATGTGGGCCAGGGCTCATGGCTGGAT 1680
Db 1729 GAAAGTTCTCAGTTCCTTCCAGGCTTCTAGAAGCATGTGGGCCAGGGCTCATGGCTGGAT 1788
Qy 1681 AATTTCCTTAGGCTTAAACAACCAAGCAAGCTTCCGCTCTCTTATTTTGGTTAAA 1740
Db 1789 AATTTCCTTAGGCTTAAACAACCAAGCAAGCTTCCGCTCTCTTATTTTGGTTAAA 1848
Qy 1741 CTTATGAATAATGATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
Db 1849 CTTATGAATAATGATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1908
Qy 1801 AGACCCAGATCACAAAGCCAAACCATGCCAGCCCTCCAGCACCCCCAGCCCCACGAC 1860
Db 1909 AGACCCAGATCACAAAGCCAAACCATGCCAGCCCTCCAGCACCCCCAGCCCCACGAC 1968
Qy 1861 CATGTTCTGAATTCGTAGACACACCGTGAGCCTTGTGACTTTAACTCATGGAAG 1920
Db 1969 CATGTTCTGAATTCGTAGACACACCGTGAGCCTTGTGACTTTAACTCATGGAAG 2028
Qy 1921 ATAACTACCTTACGTTTGAATAATGTTTCTGTGAATG 1964
Db 2029 ATAACTACCTTACGTTTGAATAATGTTTCTGTGAATG 2072

RESULT 13
ABZ11809
ID ABZ11809 standard, cDNA, 2072 BP.
XX
AC ABZ11809;
XX
DT 20-JAN-2003 (first entry)
XX
DE Human polynucleotide SEQ ID NO 691.
XX
KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
antiParkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
antiarthritic; gene; ss.
OS Homo sapiens.
XX
PN WO200270539-A2.

XX 12-SEP-2002.
XX PD
XX 05-MAR-2002; 2002WO-US005095.
XX PF
XX 05-MAR-2001; 2001US-00799451.
XX PR
XX (HYSE-) HYSEQ INC.
XX PA
XX PI Tang YT, Zhou P, Goodrich RW, Abundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2002-759812/82.
DR P-RSDB; ABP69592.
XX
PT New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.
XX
XX Claim 1; SEQ ID NO 691; 1012bp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (AB211119-
CC AB212066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;
Query Match 92.3%; Score 1812.4; DB 6; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
QY 1 ATGACGCTGTAGAGGTGCGGGCAGTTGGTGCCACAAGAGCAGAGACATCATATG 60
Db 163 ATGACGTGTAGAGGTGCGGGCAGTTGGTGCCACAAGAGCAGAGACATCATATG 222
QY 61 AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGCACAAGGGCCCTAAGCCTTTTCCA 120
Db 223 AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGCACAAGGGCCCTAAGCCTTTTCCA 282
QY 121 AGCTACAACAACAAGCTGATCATTTGGGGATTGTACATGAGACGAGCTGCCCTCTCTG 180
Db 283 AGCTACAACAACAAGCTGATCATTTGGGGATTGTACATGAGACGAGCTGCCCTCTCTG 342
QY 181 ACTGCGCGGAGCGCAAGCAAAATTCGGCGGAGATCAGCCGAAAGACCAAGTGGTGAGT 240
Db 343 ACTGCGCGGAGCGCAAGCAAAATTCGGCGGAGATCAGCCGAAAGACCAAGTGGTGAGT 402
QY 241 ATGCTGGAGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 300
Db 403 ATGCTGGAGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 462
QY 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACATTGAGAAATG 360
Db 463 GGAATGCCATGAACATCCGGGGCCCGATGTGTCACTCCTCTGAACACTGAGAAATG 522
QY 361 AAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 420
Db 523 AAGTTGAAAAACCCCGGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 582

QY 421 CACATCCAGCGCATCGACCCGGGACGTAAAGCGGACATTAAAGAGCATATATTCTTCAGG 480
Db 583 CACATCCAGCGCATCGACCCGGGACGTAAAGCGGACATTAAAGAGCATATATTCTTCAGG 642
QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTCTGGCATATGAGAGTAT 540
Db 643 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTCTGGCATATGAGAGTAT 702
QY 541 AACCCGGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 600
Db 703 AACCCGGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 762
QY 601 CTTCTGAGGAGGATGCAATCTGGGCACTGTGCACTGCTGGCCAGTGAGAGCACTCC 660
Db 763 CTTCTGAGGAGGATGCAATCTGGGCACTGTGCACTGCTGGCCAGTGAGAGCACTCC 822
QY 661 CTGACGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCACAAGAG 720
Db 823 CTGACGGATTTCACAGCCCAATGGCGGACCGTCCAGGGGCTCCAAAGCACAAGAG 882
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 780
Db 883 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCA----- 923
QY 781 GGGCAGTGTCCCGCTAGGCTGCTCATCCGATATTGATGACGGGATCTCTCCGG 840
Db 924 -----GATCTCTCCGG 936
QY 841 CTCACCTGCGCCTGTGGGACGTGTATCTGTAGAAAGCGAACAGGCGTTGATGCCGATA 900
Db 937 CTCACCTGCGCCTGTGGGACGTGTATCTGTAGAAAGCGAACAGGCGTTGATGCCGATA 996
QY 901 ACAAGATCGCCTTTAAGTTTCAGCAGAAAGCGCTCAAGAAAGCTCCAGGTGAGCCCG 960
Db 997 ACAAGATCGCCTTTAAGTTTCAGCAGAAAGCGCTCAAGAAAGCTCCAGGTGAGCCCG 1056
QY 961 TGGCAGTTTTTTGCAACCGGTTGTTGATACCTGGGCCAGGATGAGAGCACTGTGCTC 1020
Db 1057 TGGCAGTTTTTTGCAACCGGTTGTTGATACCTGGGCCAGGATGAGAGCACTGTGCTC 1116
QY 1021 AAGCATTTAAGGCGCTCTATGAAGAACTAACAAGAAAGAGGGGACCTGCCACCCCCA 1080
Db 1117 AAGCATTTAAGGCGCTCTATGAAGAACTAACAAGAAAGAGGGGACCTGCCACCCCCA 1176
QY 1081 GCCAAACCCGAGCAAGGGTCTGTGGCATCCAGGCGCTGTGCCGCTTCACTGACGGGAG 1140
Db 1177 GCCAAACCCGAGCAAGGGTCTGTGGCATCCAGGCGCTGTGCCGCTTCACTGACGGGAG 1236
QY 1141 ACCCTGCAAGGGGAGCAGGACGCGCTCCAGGCCACCAAGCCCGGTTCCCGGGCCC 1200
Db 1237 ACCCTGCAAGGGGAGCAGGACGCGCTCCAGGCCACCAAGCCCGGTTCCCGGGCCC 1296
QY 1201 ATTTGTCAGCTTCCCGGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGTGGGCT 1260
Db 1297 ATTTGTCAGCTTCCCGGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGTGGGCT 1356
QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAG 1320
Db 1357 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAG 1416
QY 1321 GGAAGACCTCAGGGTCTCTGAGATTCTCTGACGTGGAATCCATGCCCCGCTCCCAAG 1380
Db 1417 GGAAGACCTCAGGGTCTCTGAGATTCTCTGACGTGGAATCCATGCCCCGCTCCCAAG 1476
QY 1381 GACCTGAGCGTAGAGGGCCCTTGGTTCGCCCATTAATTTACAGACAGAGCTGTGGGTC 1440
Db 1477 GACCTGAGCGTAGAGGGCCCTTGGTTCGCCCATTAATTTACAGACAGAGCTGTGGGTC 1536
QY 1441 CGTGCCATATCCAGAGAGCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1500
Db 1537 CGTGCCATATCCAGAGAGCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1596

QY 1501 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCACCCCCTTCAGA 1560
DB 1597 CGGGTAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGCACCCCCTTCAGA 1656
QY 1561 GCTAGGACGAACAGCCGTGCTGCCACCTCAGGGCCTTGCTGCGGCTCCACTTG 1620
DB 1657 GCTAGGACGAACAGCCGTGCTGCCACCTCAGGGCCTTGCTGCGGCTCCACTTG 1716
QY 1621 GAAAGTTCAGTTCCTCCCTCCAGGCTTCTAGAAGCATCTGGGCCAAGGCTCATGGCTGAT 1680
DB 1717 GAAAGTTCAGTTCCTCCCTCCAGGCTTCTAGAAGCATCTGGGCCAAGGCTCATGGCTGAT 1776
QY 1681 AATTTCCTAGGCTTAACAACCCAGCAAGCTTCGCGTCCCTGTTTATTGTTAA 1740
DB 1777 AATTTCCTAGGCTTAACAACCCAGCAAGCTTCGCGTCCCTGTTTATTGTTAA 1836
QY 1741 CTTATGAAATGTATTAAAGAAAGTGCAGCTCGAGAGATTCAAGATGAAACACACC 1800
DB 1837 CTTATGAAATGTATTAAAGAAAGTGCAGCTCGAGAGATTCAAGATGAAACACACC 1896
QY 1801 AGACCCCAAGATCAAAAAGCCAAACCATGCCCCCTCCAGCAACCCCAAGCCCAAGAC 1860
DB 1897 AGACCCCAAGATCAAAAAGCCAAACCATGCCCCCTCCAGCAACCCCAAGCCCAAGAC 1956
QY 1861 CATCGTTCGAATTCTGACGACACCCGTGACCTGCTTGTACTTAACTCATGAAAG 1920
DB 1957 CATCGTTCGAATTCTGACGACACCCGTGACCTGCTTGTACTTAACTCATGAAAG 2016
QY 1921 ATAACCTACCTTCACGTTTGAATAATGTTTCTGTTGAAATG 1964
DB 2017 ATAACCTACCTTCACGTTTGAATAATGTTTCTGTTGAAATG 2060

RESULT 14

ADM44327
ID ADM44327 standard; cDNA; 2072 BP.

AC ADM44327;

DT 03-JUN-2004 (first entry)

DE Novel human arginine-rich protein cDNA #691.

KW ss; gene; human; arginine-rich protein; cancer; inflammation;
genetic disorder.

OS Homo sapiens.

PN US2004053250-A1.

PD 18-MAR-2004.

PF 21-NOV-2002; 2002US-00302172.

PR 05-MAR-2001; 2001US-00799451.

PR 05-MAR-2002; 2002WO-US005095.

PR 20-AUG-2002; 2002US-00225251.

PA (TANG/) TANG Y T.

PA (XUE/) XUE A.

PA (DRMA/) DRMANAC R T.

PI Tang YT, Xue A, Drmanac RT;

DR WPI; 2004-238579/22.

PT New isolated arginine-rich protein-like polynucleotides and polypeptides,
useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.

PS Disclosure; SEQ ID NO 691; 51pp; English.

XX

CC The invention relates to an isolated polynucleotide. The methods and
CC compositions of the present invention are useful for the diagnosis and/or
CC treatment of diseases or conditions associated with aberrant expression
CC or activity of the arginine-rich protein-like polypeptides, such as
CC cancer and inflammation. They can also be used in forensics, gene
CC mapping, identification of mutations responsible for genetic disorders,
CC and in assessing biodiversity. The present sequence represents a novel
CC human arginine-rich protein cDNA.
SQ Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;

Query Match 92.3%; Score 1812.4; DB 12; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY 1 ATGACCTGTAGAGGTCGCGGCACTTGTGGGCACAAGCCAGAGACATCATATG 60
DB 163 ATGACCTGTAGAGGTCGCGGCACTTGTGGGCACAAGCCAGAGACATCATATG 222
QY 61 AAATACGAAAAGGACACCCGAGCTGGGCTCCAGAGCAAGGGCTTAAGCCTTTTGA 120
DB 223 AAATACGAAAAGGACACCCGAGCTGGGCTCCAGAGCAAGGGCTTAAGCCTTTTGA 282
QY 121 AGCTACAACAACAAGCTGATCATTTGGGATTGTACATAGACGAGCTGCTCTG 180
DB 283 AGCTACAACAACAAGCTGATCATTTGGGATTGTACATAGACGAGCTGCTCTG 342
QY 181 ACTGCGGGAAGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 240
DB 343 ACTGCGGGAAGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 402
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGTCGAGCGTACAG 300
DB 403 ATGCTGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGTCGAGCGTACAG 462
QY 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATG 360
DB 463 GGAATGCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATG 522
QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCTATGAG 420
DB 523 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCTATGAG 582
QY 421 CACATCCAGCGCATCGACCCGGAAGCTAAGCGGACATTAAGGAACATATATTCTCAGG 480
DB 583 CACATCCAGCGCATCGACCCGGAAGCTAAGCGGACATTAAGGAACATATATTCTCAGG 642
QY 481 GATCGATACGGAACCAAGCAGCCGGAATCTACTCCAATCTCTGGCATATGAGAGTAT 540
DB 643 GATCGATACGGAACCAAGCAGCCGGAATCTACTCCAATCTCTGGCATATGAGAGTAT 702
QY 541 AACCCGAGGTGGGCTACTGACAGGAACTGAGCCACATGCGCGCTTGTCTCTCTAT 600
DB 703 AACCCGAGGTGGGCTACTGACAGGAACTGAGCCACATGCGCGCTTGTCTCTCTAT 762
QY 601 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTCAGTCTGCGCAGTGAAGGCACTCC 660
DB 763 CTTCTGAGAGAGATGATTTCTGGGCACTGTGTCAGTCTGCGCAGTGAAGGCACTCC 822
QY 661 CTGACGGATTTACAGCCCAATGGCGGAGCTCCAGGGGCTCCAAGACCAAGAG 720
DB 823 CTGACGGATTTACAGCCCAATGGCGGAGCTCCAGGGGCTCCAAGACCAAGAG 882
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAGAAAGATCTATGT 780
DB 883 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCA----- 923
QY 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
DB 924 -----GATCTCTCTCGGG 936
QY 841 CTCACCTGCGCCTGTGGGAGCTGTATCTGTAGAAAGCGAAACAGGCGTTGATGCCGATA 900

Db	937	CTCACCCCTGCGCCTGTGTGGGACGTGTATCTGGTAGAAGGCCAACAGGCGTTGATGCCCCATA	996
Qy	901	ACAAGAATCGCCTTTAAGGTTACGACAGAAGCGCCTCACGAAGACGTCAGAGTGTGGCCCG	960
Db	997	ACAAGAATCGCCTTTAAGGTTACGACAGAAGCGCCTCACGAAGACGTCAGAGTGTGGCCCG	1056
Qy	961	TGGGCACGTTTTTGGCAACCGGTCGTTGATACCTGGGCCAGGGATGAGGACACTGTGCTC	1020
Db	1057	TGGGCACGTTTTTGGCAACCGGTCGTTGATACCTGGGCCAGGGATGAGGACACTGTGCTC	1116
Qy	1021	AAGCATCTTAGGGCCTCTATGAAAGAACTAACAAAGAAAGCAGGGGACCTGCAACCCCA	1080
Db	1117	AAGCATCTTAGGGCCTCTATGAAAGAACTAACAAAGAAAGCAGGGGACCTGCAACCCCA	1176
Qy	1081	GCCAAACCCGAGCAAGGTCGTCGGCATCCAGGCGCTGTGCCGCTTCACGTGCGGGAAG	1140
Db	1177	GCCAAACCCGAGCAAGGTCGTCGGCATCCAGGCGCTGTGCCGCTTCACGTGCGGGAAG	1236
Qy	1141	ACCCCTGCAAGGGGGACAGGCAAGGCCCTCCAGGCCACCAAGCCCGTTCCCGCGGCC	1200
Db	1237	ACCCCTGCAAGGGGGACAGGCAAGGCCCTCCAGGCCACCAAGCCCGTTCCCGCGGCC	1296
Qy	1201	ATTGGTCAGCTTCCCCGCCACGGGCACTCGTTCCTTCACACCTGTCTGTGGGCT	1260
Db	1297	ATTGGTCAGCTTCCCCGCCACGGGCACTCGTTCCTTCACACCTGTCTGTGGGCT	1356
Qy	1261	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG	1320
Db	1357	GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTAG	1416
Qy	1321	GGAGGACCTCAGGGTTCCTGGAGATTCTGCAGTGGAACTCCATGCCCGCTCCCAAG	1380
Db	1417	GGAGGACCTCAGGGTTCCTGGAGATTCTGCAGTGGAACTCCATGCCCGCTCCCAAG	1476
Qy	1381	GACCTGACGTAGAGGGCCCTTGGTTCGCCATTATGATTTCAACAAGAGCTGTGGTTC	1440
Db	1477	GACCTGACGTAGAGGGCCCTTGGTTCGCCATTATGATTTCAACAAGAGCTGTGGTTC	1536
Qy	1441	CGTGCCATATCCCAAGAGGACCAAGCTGGCCCCCTGCTGGCAGGGCTGAACAACCTTGGGAG	1500
Db	1537	CGTGCCATATCCCAAGAGGACCAAGCTGGCCCCCTGCTGGCAGGGCTGAACAACCTTGGGAG	1596
Qy	1501	CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGGCACCCCTTCAGA	1560
Db	1597	CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGGCACCCCTTCAGA	1656
Qy	1561	GCTAGGAGCAACAGCCGTCGTCTCCACCTCAGGGCCTTGCTTGCGGCTCCACTTG	1620
Db	1657	GCTAGGAGCAACAGCCGTCGTCTCCACCTCAGGGCCTTGCTTGCGGCTCCACTTG	1716
Qy	1621	GAAAGTTCACGTTCCCTCCAGGCTTCTAGAAGCATGTGGGCCAGGGCTCATGGCTGGAT	1680
Db	1717	GAAAGTTCACGTTCCCTCCAGGCTTCTAGAAGCATGTGGGCCAGGGCTCATGGCTGGAT	1776
Qy	1681	AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCCGCTCTCGTTTATTTTGGTTAAA	1740
Db	1777	AATTTCCCTAGGCTTAAACAACCAAGCAAGCTTCCGCTCTCGTTTATTTTGGTTAAA	1836
Qy	1741	CTTATGAAAATGTATTAGAAAAGAGTGACGCTCGAGAGAGATTCAAGATGGAACACACC	1800
Db	1837	CTTATGAAAATGTATTAGAAAAGAGTGACGCTCGAGAGAGATTCAAGATGGAACACACC	1896
Qy	1801	AGACCCGAGATCAAAAGCCAACCATGCCCCAGGCCCTCCAGCACCCCCAGCCCAAGAC	1860
Db	1897	AGACCCGAGATCAAAAGCCAACCATGCCCCAGGCCCTCCAGCACCCCCAGCCCAAGAC	1956
Qy	1861	CATCGTTCTGAATTCTGACGACACCCGTGAGCCTGCCTTTGACTTTAAACTCATGAAGG	1920
Db	1957	CATCGTTCTGAATTCTGACGACACCCGTGAGCCTGCCTTTGACTTTAAACTCATGAAGG	2016
Qy	1921	ATAACTACCTTCACGTTTGAAAATAAATGTTTCTGTGAAATG 1964	
Db	2017	ATAACCACTTCATGTGTTTGAAAATAAATGTTTCTGTGAAATG 2060	

```

RESULT 15
ADC37388
ID      ADC37388 standard; DNA; 2647 BP.
XX
AC      ADC37388;
XX
XX      18-DEC-2003 (first entry)
DT
XX      Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 221.
DE
XX      Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
XX      cancer; infectious disease; bone disease; AIDS;
XX      neurodegenerative disease; ischaemic disorder; Antinflammatory;
XX      Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
XX      Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
OS      Homo sapiens.
XX
XX      WO2003048202-A2.
XX
XX      12-JUN-2003.
PD
XX      03-DEC-2002; 2002WO-JP012644.
PF
XX      03-DEC-2001; 2001JP-00368692.
PR      05-DEC-2001; 2001US-0335829P.
PR      03-OCT-2002; 2002JP-00291302.
PR      04-OCT-2002; 2002US-0415769P.
XX
XX      (ASAH ) ASAH KASEI KK.
XX
XX      Matsuda A, Muramatsu S;
XX
XX      WPI; 2003-505282/47.
DR      P-PSDB; ADC37389.
XX
XX      New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT      useful for treating inflammation, autoimmune diseases, cancers,
PT      infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT      ischemic disorders.
XX
XX      Claim 4; SEQ ID NO 221; 938bp; English.
XX
XX      The present invention relates to novel proteins and their coding
CC      sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC      kappaB). The proteins and their coding sequences are useful for treating
CC      a disease associated with NF-kappaB activation, such as inflammation,
CC      autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC      neurodegenerative diseases, or ischaemic disorders.
XX
XX      Sequence 2647 BP; 608 A; 766 C; 750 G; 523 T; 0 U; 0 Other;
SQ

Query Match          92.1%; Score 1809; DB 10; Length 2647;
Best local Similarity 96.4%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 5; Indels 66; Gaps 1;

QY      1 ATGACGCTGTAGAGTCCGCGGCACTTGTGGCAACAAGACGAGACATCATTTATG 60
DB      755 ATGACGTGTAGAGTCCGCGGCACTTGTGGCAACAAGACGAGACATCATTTATG 814
QY      61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 120
DB      815 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 874
QY      121 AGCTACAACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCCTCTG 180
DB      875 AGCTACAACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCCTCTG 934
QY      181 ACTGCGCGGAGCGCAAGCAATTCCGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT 240
DB      935 ACTGCGCGGAGCGCAAGCAATTCCGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT 994

```


QY 241 ATGCTGGAGACTGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 300
 Db 995 ATGCTGGAGAGACTGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG 1054
 QY 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAAATG 360
 Db 1055 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGGAAATG 1114
 QY 361 AAGTTGAAAAAACCCTGGAGATACAGATCATGAAGGAAAGGCAAGAGTCACTGAG 420
 Db 1115 AAGTTGAAAAAACCCTGGAGATACAGATCATGAAGGAAAGGCAAGAGTCACTGAG 1174
 QY 421 CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAGGAAGCATATATTCTTCAGG 480
 Db 1175 CACATCCAGCGCATCGACCGGAGCTAAGCGGACATTAAGGAAGCATATATTCTTCAGG 1234
 QY 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTGCGATATGAGAGTAT 540
 Db 1235 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTGCGATATGAGAGTAT 1294
 QY 541 AACCCTGGAGGTGGGCTACTGACAGGGAACCTGAGCCACATCGCCCTTGTCTCTCTAT 600
 Db 1295 AACCCTGGAGGTGGGCTACTGACAGGGAACCTGAGCCACATCGCCCTTGTCTCTCTAT 1354
 QY 601 CTTCTGAGGAGATGCAATCTGAGGCACTGTGCACTGTGCGCAGTGAAGAGCACTCC 660
 Db 1355 CTTCTGAGGAGATGCAATCTGAGGCACTGTGCACTGTGCGCAGTGAAGAGCACTCC 1414
 QY 661 CTGCAAGGATTTACAGCCCAATGCGGAGCCGTCCAGGGGCTCCAAGCAACAGAGAG 720
 Db 1415 CTGCAAGGATTTACAGCCCAATGCGGAGCCGTCCAGGGGCTCCAAGCAACAGAGAG 1474
 QY 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGCAATCAGACAAAGAAATCTATGT 780
 Db 1475 CATGTGTAGCCACGTCACAAACCAAGACCATGGGCAATCA----- 1515
 QY 781 GGGCAGTGTCCCCCGTTAGCGTCCCTCATCCGATATTGATGACGGATCTCTCGGG 840
 Db 1516 -----GATCTCTCTCGGG 1528
 QY 841 CTCACCCCTGCGCTGTGGAGCGTGTATCTGTGTAAGGCGAAACAGCGTGTATGCCGATA 900
 Db 1529 CTCACCCCTGCGCTGTGGAGCGTGTATCTGTGTAAGGCGAAACAGCGTGTATGCCGATA 1588
 QY 901 ACAAGATCGCCTTTAAGTTACAGAGAAGCGCTCAAGAGACGTCCAGGTGTGCGCG 960
 Db 1589 ACAAGATCGCCTTTAAGTTACAGAGAAGCGCTCAAGAGACGTCCAGGTGTGCGCG 1648
 QY 961 TGGCAGCGTTTTCACACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1020
 Db 1649 TGGCAGCGTTTTCACACCGGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1708
 QY 1021 AAGCATCTTAGGGCCTTATGAAGAACTAACAAAGCAAGGGGACCTGCCACCCCA 1080
 Db 1709 AAGCATCTTAGGGCCTTATGAAGAACTAACAAAGCAAGGGGACCTGCCACCCCA 1768
 QY 1081 GCCAAACCCGAGCAAGGCTGTGGCATCCAGGCTGTGCGGCTTCACTGTGGCGGAAG 1140
 Db 1769 GCCAAACCCGAGCAAGGCTGTGGCATCCAGGCTGTGCGGCTTCACTGTGGCGGAAG 1828
 QY 1141 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCAAGGCCCAACAGCCGGTTCCCGCGGCC 1200
 Db 1829 ACCCTCTGCAAGGGGAGCAGGAGGCGCCCTCAAGGCCCAACAGCCGGTTCCCGCGGCC 1888
 QY 1201 ATTGGTCAAGCTTCCCGCAAGGACCTGTTCTTCCACACCTGTCTGTGGGCT 1260
 Db 1889 ATTGGTCAAGCTTCCCGCAAGGACCTGTTCTTCCACACCTGTCTGTGGGCT 1948
 QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGCTGTCCAGGCCCGGCTGGCTCAG 1320
 Db 1949 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGCTGTCCAGGCCCGGCTGGCTCAG 2008

QY 1321 GAGGACCTCAGGGTTCCTGAGAGATTCTGCACTGGAATCCATGCCCCGCTCCCAACG 1380
 Db 2009 GAGGACCTCAGGGTTCCTGAGAGATTCTGCACTGGAATCCATGCCCCGCTCCCAACG 2068
 QY 1381 GACTGGAAGTAGAGGGCCCTTGTGTTCCGCCATTAATGATTTCAGACAGAGCTGTGGTC 1440
 Db 2069 GACTGGAAGTAGAGGGCCCTTGTGTTCCGCCATTAATGATTTCAGACAGAGCTGTGGTC 2128
 QY 1441 CGTGCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGCAAGGCTGAACACCTGCGAG 1500
 Db 2129 CGTGCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGCAAGGCTGAACACCTGCGAG 2188
 QY 1501 CGGTTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCCGACAGGGCAACCTTCAGA 1560
 Db 2189 CGGTTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCCGACAGGGCAACCTTCAGA 2248
 QY 1561 GCTAGGAGCAAGACCGGTGTGCTCCACCTCAAGGCTTGCCCTGCGGCTCCACTTG 1620
 Db 2249 GCTAGGAGCAAGACCGGTGTGCTCCACCTCAAGGCTTGCCCTGCGGCTCCACTTG 2308
 QY 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
 Db 2309 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAGCATCTGGGCCAGGGCTCATGGCTGAT 2368
 QY 1681 AATTTCCTAGGCTTACAAACCCAGCAAGCTTCCGCTCTCGTTTATTTTGTAA 1740
 Db 2369 AATTTCCTAGGCTTACAAACCCAGCAAGCTTCCGCTCTCGTTTATTTTGTAA 2428
 QY 1741 CTTATGAAATGTATTAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
 Db 2429 CTTATGAAATGTATTAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 2488
 QY 1801 AGACCCAGATCAAAAGCCAAACCATGCCCCCTCCAGACACCCCAAGCCAGAC 1860
 Db 2489 AGACCCAGATCAAAAGCCAAACCATGCCCCCTCCAGACACCCCAAGCCAGAC 2548
 QY 1861 CATGTTCTGAATTCTGACGACACCGTGAGCCTTGTACTTAACTCATGAAGG 1920
 Db 2549 CATGTTCTGAATTCTGACGACACCGTGAGCCTTGTACTTAACTCATGAAGG 2608
 QY 1921 ATAACTACCTTCAAGTTTGAATAATGTTTCTGTG 1959
 Db 2609 ATAACTACCTTCAAGTTTGAATAATGTTTCTGTG 2647

Search completed: April 5, 2006, 13:32:41
 Job time : 821 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:54:55 ; Search time 362 Seconds
(without alignments)
9644.009 Million cell updates/sec

Title: US-10-071-838-1
Perfect score: 1964
Sequence: 1 atggacgtgtagagtcgc.....aatgttcctgtgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/1_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/H_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq: *
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq: *
9: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1836.4	93.5	2084	US-09-799-451-692	Sequence 692, App
2	1812.4	92.3	2072	US-09-799-451-691	Sequence 691, App
3	999	50.9	8201	US-08-253-155A-9	Sequence 9, Appli
4	121	6.2	121	US-09-513-999C-24808	Sequence 24808, A
5	92	4.7	1925	US-09-949-016-2139	Sequence 2139, Ap
6	74.2	3.8	187	US-09-513-999C-13267	Sequence 13267, A
7	74	3.8	6453	US-08-306-691B-14	Sequence 14, Appl
8	74	3.8	6453	US-09-209-668-10	Sequence 10, Appl
9	74	3.8	6453	US-09-356-952-8	Sequence 8, Appli
10	72.6	3.7	1981	US-09-620-312D-715	Sequence 715, App
11	70	3.6	1823	US-09-774-528-336	Sequence 336, App
12	70	3.6	1823	US-10-120-988-336	Sequence 336, App
13	68.4	3.5	7218	US-08-232-463-14	Sequence 14, Appl
14	63.4	3.2	806	US-09-270-767-9899	Sequence 9899, Ap
15	53.6	2.7	4039	US-08-363-300-1	Sequence 1, Appli
16	52.4	2.7	69199	US-09-949-016-13881	Sequence 13881, A
17	46.8	2.4	1887	US-09-270-767-12482	Sequence 12482, A
18	45.4	2.3	601	US-09-949-016-51412	Sequence 51412, A
19	45.4	2.3	25431	US-09-949-016-13234	Sequence 13234, A
20	45	2.3	2066	US-10-104-047-56	Sequence 56, Appl
21	42.6	2.2	2955	US-09-620-312D-676	Sequence 676, App
22	42.6	2.2	3039	US-09-620-312D-675	Sequence 675, App
23	42	2.1	289	US-09-007-005-17	Sequence 17, Appl
24	42	2.1	289	US-09-244-796-17	Sequence 17, Appl

25	41.4	2.1	505	3	US-09-621-976-15639	Sequence 15639, A
26	40.4	2.1	1984	3	US-09-949-016-3047	Sequence 3047, Ap
27	40.2	2.0	13987	2	US-08-804-227C-13	Sequence 13, Appl
28	40.2	2.0	44377	2	US-08-804-227C-7	Sequence 7, Appli
29	40.2	2.0	53799	3	US-10-042-665A-3	Sequence 3, Appli
30	40.2	2.0	723	3	US-09-270-767-1491	Sequence 1491, Ap
31	39.8	2.0	723	3	US-09-270-767-1491	Sequence 1491, Ap
32	39.8	2.0	723	3	US-09-270-767-1491	Sequence 1491, Ap
33	39.8	2.0	33478	3	US-09-949-016-13615	Sequence 13615, A
34	39	2.0	505	3	US-09-621-976-15639	Sequence 15639, A
35	38.6	2.0	482	3	US-09-392-184-8	Sequence 8, Appli
36	38.6	2.0	15722	3	US-09-949-016-16709	Sequence 16709, A
37	38.6	2.0	17000	3	US-09-548-797B-7	Sequence 7, Appli
38	38.6	2.0	17425	3	US-09-511-625B-5	Sequence 5, Appli
39	38.2	1.9	3983	3	US-09-762-311-3	Sequence 3, Appli
40	38.2	1.9	3988	3	US-09-762-311-4	Sequence 4, Appli
41	38	1.9	1179	3	US-10-077-176-71	Sequence 71, Appl
42	37.8	1.9	97423	3	US-09-949-016-12742	Sequence 12742, A
43	37.8	1.9	97424	3	US-09-949-016-15576	Sequence 15576, A
44	37.6	1.9	1056	3	US-09-902-540-5688	Sequence 5688, Ap
45	37.6	1.9	1152	3	US-09-902-540-2101	Sequence 2101, Ap

ALIGNMENTS

RESULT 1
US-09-799-451-692
Sequence 692, Application US/09799451
Patent No. 6783969
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyang
APPLICANT: Zhang, Jie
APPLICANT: Xue, Aidong J.
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ma, Yungqing
APPLICANT: Yamazaki, Victoria
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Zhiwei
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163) .. (1755)
US-09-799-451-692

Query Match 93.5%; Score 1836.4; DB 3; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

Qy 1 ATGACGTGCTAGAGGTGCGGCGCAGTTGTGGGACAGAGCGAGAGCATTCATTATG 60
Db 163 ATGACGTGCTAGAGGTGCGGCGCAGTTGTGGGACAGAGCGAGAGCATTCATTATG 222

QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGCGCTAAGCCTTTTCGA 120
| | | | |
Db 223 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGCGCTAAGCCTTTTCGA 282
QY 121 AGCTACAAACAACAGTTCGATCATTTGGGGATTGTACATGAGACGAGCTGCCCTCTTG 180
| | | | |
Db 283 AGCTACAAACAACAGTTCGATCATTTGGGGATTGTACATGAGACGAGCTGCCCTCTTG 342
QY 181 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240
| | | | |
Db 343 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGAT 402
QY 241 ATGCTGGAGACTGGGAGAATACAAAAGCAGAGAAAGCTCATAGATCGAGCGTACAG 300
| | | | |
Db 403 ATGCTGGAGACTGGGAGAATACAAAAGCAGAGAAAGCTCATAGATCGAGCGTACAG 462
QY 301 GGAATGCCCATGACATCCGGGGCGCCGATGTGTCACTCCTCTGAACATTGAGGAATG 360
| | | | |
Db 463 GGAATGCCCATGACATCCGGGGCGCCGATGTGTCACTCCTCTGAACACTGAGGAATG 522
QY 361 AAGTTGAAAAAAGCCCGGAAGATACCATGATGTAAGAGAAAGGCGCAAGGTCACTGAG 420
| | | | |
Db 523 AAGTTGA----- 529
QY 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAAGAAACATATATTCTTCAG 480
| | | | |
Db 530 -ACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAAGAAACATATATTCTTCAG 588
QY 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGCGATATGAGAGATAT 540
| | | | |
Db 589 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGCGATATGAGAGATAT 648
QY 541 AACCCGAGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
| | | | |
Db 649 AACCCGAGTGGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 708
QY 601 CTTCTGAGAGAGATGCATTTCTGGGCACTGTGCAAGTGTCTGGCCAGTGAAGGCACTCC 660
| | | | |
Db 709 CTTCTGAGAGAGATGCATTTCTGGGCACTGTGCAAGTGTCTGGCCAGTGAAGGCACTCC 768
QY 661 CTGCAAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 720
| | | | |
Db 769 CTGCAAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAACAGAG 828
QY 721 CATGTGTAGCCAAGTCAACAACCAAGACCATGGGCACTCAGGACCAAGAAAGATCTATGT 780
| | | | |
Db 829 CATGTGTAGCCAAGTCAACAACCAAGACCATGGGCACTCAGGACCAAGAAAGATCTATGT 888
QY 781 GGGCAGTGTTCCTCGTTAGGCTGCCTCATCCGATATTGATTGACGGGATCTCTCGGG 840
| | | | |
Db 889 GGGCAGTGTTCCTCGTTAGGCTGCCTCATCCGATATTGATTGACGGGATCTCTCGGG 948
QY 841 CTCACCCCTGCGCTGTGGAGCGTGTATCTGGTAGAAGGCGAAGCGGCTTGATGCCGATA 900
| | | | |
Db 949 CTCACCCCTGCGCTGTGGAGCGTGTATCTGGTAGAAGGCGAAGCGGCTTGATGCCGATA 1008
QY 901 ACAAGAAATCGCCTTTAAGGTTCAAGCAGAACGCGCTCACGAAGACGTCCAGGTGTGCCG 960
| | | | |
Db 1009 ACAAGAAATCGCCTTTAAGGTTCAAGCAGAACGCGCTCACGAAGACGTCCAGGTGTGCCG 1068
QY 961 TGGCAGCGTTTTTGCAACCGGTTCTGTGATACCTGGGCGCAGGGATGAGACACTGTGCTC 1020
| | | | |
Db 1069 TGGCAGCGTTTTTGCAACCGGTTCTGTGATACCTGGGCGCAGGGATGAGACACTGTGCTC 1128
QY 1021 AAGCATCTTAGGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
| | | | |
Db 1129 AAGCATCTTAGGGCTCTATGAAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1188
QY 1081 GCCAAACCCGAGCAAGGGTGTCTGGCATCCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1140
| | | | |
Db 1189 GCCAAACCCGAGCAAGGGTGTCTGGCATCCAGGCTGTGCGGCTTCAAGTGGCGGGAAG 1248
QY 1141 ACCCTCTGCAAGGGGGAAGGCAAGGCGCCCTCCAGGCCACCAAGCCCGGTTCCCGCGGCC 1200

Db 1249 ACCCTCTGCAAGGGGGAAGAGGACAGGCGCCCTCCAGGCCCAAGCCGGTTCCCGGGCCC 1308
QY 1201 ATTTGTGACGTTCCCGGCAAGGCAAGCTGTTCTTCCACACCCCTGTCTGTGGGGCT 1260
| | | | |
Db 1309 ATTTGTGACGTTCCCGGCAAGGCAAGCTGTTCTTCCACACCCCTGTCTGTGGGGCT 1368
QY 1261 GTCCGGGAAGACACCTACCCTGTGGCACTCAGGGTGTGCCAGCCCGGCGCTGCTCAG 1320
| | | | |
Db 1369 GTCCGGGAAGACACCTACCCTGTGGCACTCAGGGTGTGCCAGCCCGGCGCTGCTCAG 1428
QY 1321 GAGGACCTCAGGGTTCCTGGAGATTCTGCAGTGGAACCTCATGCCCCGCTCCCAAG 1380
| | | | |
Db 1429 GAGGACCTCAGGGTTCCTGGAGATTCTGCAGTGGAACCTCATGCCCCGCTCCCAAG 1488
QY 1381 GACCTGACGTAGAGGGCCCTTGGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1440
| | | | |
Db 1489 GACCTGACGTAGAGGGCCCTTGGTTCCGCCATTATGATTTCAGACAGAGCTGTGGGTC 1548
QY 1441 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCCCTGGAG 1500
| | | | |
Db 1549 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCCCTGGAG 1608
QY 1501 CGGCTGACATCGGCTTTCGCTGCAACCCAGACATGATTCGACCAAGGCAACCCCTTCAGA 1560
| | | | |
Db 1609 CGGCTGACATCGGCTTTCGCTGCAACCCAGACATGATTCGACCAAGGCAACCCCTTCAGA 1668
QY 1561 GCTAGGACGAACAAGCCGTGTGCTCCCACTCAGGGGCTTGCTCTGCGGCTTCCACTTG 1620
| | | | |
Db 1669 GCTAGGACGAACAAGCCGTGTGCTCCCACTCAGGGGCTTGCTCTGCGGCTTCCACTTG 1728
QY 1621 GAAAGTTCTAGTTCCTCCAGGCTTCTAGAAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
| | | | |
Db 1729 GAAAGTTCTAGTTCCTCCAGGCTTCTAGAAAGCATCTGGGCCAGGGCTCATGGCTGAT 1788
QY 1681 AATTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCGTCTCGTTTATTGTTGTTAA 1740
| | | | |
Db 1789 AATTCCCTAGGCTTAAACAACCAAGCAAGCTTCGCGTCTCGTTTATTGTTGTTAA 1848
QY 1741 CTTATGAATATGTATTAAGAAAGTGCAGCTGCAGAGAGATTCAAGATGGAACACACC 1800
| | | | |
Db 1849 CTTATGAATATGTATTAAGAAAGTGCAGCTGCAGAGAGATTCAAGATGGAACACACC 1908
QY 1801 AGACCCCAAGATCAAAAGCCAAACCATGCCAGGCCCTCCCAAGACACCCCAAGCAC 1860
| | | | |
Db 1909 AGACCCCAAGATCAAAAGCCAAACCATGCCAGGCCCTCCCAAGACACCCCAAGCAC 1968
QY 1861 CATGTTCTGAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTAAACTCATGGAAG 1920
| | | | |
Db 1969 CATGTTCTGAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTAAACTCATGGAAG 2028
QY 1921 ATAACTACCTTCACGTTTGAATAAATGTTTCTGTGAAATG 1964
| | | | |
Db 2029 ATAACTACCTTCATGTTTGAATAAATGTTTCTGTGAAATG 2072

RESULT 2
US-09-799-451-691
; Sequence 691, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yuning
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong

APPLICANT: Wang, Zhiwei
APPLICANT: Wang, Dunrui
APPLICANT: Yang, Yonghong
APPLICANT: Wehrman, Tom
APPLICANT: Ghosh, Reena
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6783969e1 Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 803
CURRENT APPLICATION NUMBER: US/09/799,451
CURRENT FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 948
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 691
LENGTH: 2072
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163)..(1743)
US-09-799-451-691

Query Match 92.3%; Score 1812.4; DB 3; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY 1 ATGACGTGCTAGAGTTCGGCGGCACTTGTGGGACAAAGCCGAGACATCATTTATG 60
DB 163 ATGACGTGCTAGAGTTCGGCGGCACTTGTGGGACAAAGCCGAGACATCATTTATG 222
QY 61 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA 120
DB 223 AAATACGAAAGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA 282
QY 121 AGCTACAAACAACGTCGATCACTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG 180
DB 283 AGCTACAAACAACGTCGATCACTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG 342
QY 181 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGGAGATCAAGCCGAAAGCAAGTGGGTGAT 240
DB 343 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGGAGATCAAGCCGAAAGCAAGTGGGTGAT 402
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAAAAGCTCATAGTCGAGGTAACAAG 300
DB 403 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAAAAGCTCATAGTCGAGGTAACAAG 462
QY 301 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 360
DB 463 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 522
QY 361 AAGTTGAAAAAAGGAGAAATACAGATCATGAAGGAGAAAGGCAAGAGTCACTTGAG 420
DB 523 AAGTTGAAAAAAGGAGAAATACAGATCATGAAGGAGAAAGGCAAGAGTCACTTGAG 582
QY 421 CACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGGAAGCATATATTCTCAGG 480
DB 583 CACATCCAGCGCATCGACCGGACGTAAGCGGACATTAAGGAAGCATATATTCTCAGG 642
QY 481 GATCGATACGGAACCAAGCAGGGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT 540
DB 643 GATCGATACGGAACCAAGCAGGGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT 702
QY 541 AACCCGAGGTGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
DB 703 AACCCGAGGTGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 762
QY 601 CTTCCTGAGGAGATGATCTGCGGCACTGTGACGCTGTGCGCAGTGAAGCACTCC 660
DB 763 CTTCCTGAGGAGATGATCTGCGGCACTGTGACGCTGTGCGCAGTGAAGCACTCC 822
QY 661 CTGAGGAGATTTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
DB 823 CTGAGGAGATTTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 882

QY 721 CATGTGTAGCCACGTCACAAACCCAAAGCCATGGGGCATCAGGACCAAGAAATCTATGT 780
DB 883 CATGTGTAGCCACGTCACAAACCCAAAGCCATGGGGCATCA----- 923
QY 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATGATTGACGGGATCTCTCGGG 840
DB 924 -----GATCTCTCGGG 936
QY 841 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGCGTTGATGCCGATA 900
DB 937 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGCGTTGATGCCGATA 996
QY 901 ACAAGATCGCCTTTAAGTTACGACGAAGCGCCTCAAGAAAGCTCCAGTGTGGCCG 960
DB 997 ACAAGATCGCCTTTAAGTTACGACGAAGCGCCTCAAGAAAGCTCCAGTGTGGCCG 1056
QY 961 TGGGACGTTTTTGCACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1020
DB 1057 TGGGACGTTTTTGCACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTGCTC 1116
QY 1021 AAGCATCTTAGGGCTCTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCAACCCCA 1080
DB 1117 AAGCATCTTAGGGCTCTATGAAGAACTTAACAAGAAAGCAGGGGACCTGCAACCCCA 1176
QY 1081 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCACTGTGGCGGAAG 1140
DB 1177 GCCAAACCCGAGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCACTGTGGCGGAAG 1236
QY 1141 ACCCTCTGCAAGGGGGACAGGCAAGGCCCCCTCCAGGCCCAACAGCCCCGTTCCCGGGCC 1200
DB 1237 ACCCTCTGCAAGGGGGACAGGCAAGGCCCCCTCCAGGCCCAACAGCCCCGTTCCCGGGCC 1296
QY 1201 ATTTGTCAGCTTCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT 1260
DB 1297 ATTTGTCAGCTTCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT 1356
QY 1261 GTCCGGAAAGACACCTACCTGTGGGCACTCAGGCTGTGCCAGCCCCGCTGCTCAG 1320
DB 1357 GTCCGGAAAGACACCTACCTGTGGGCACTCAGGCTGTGCCAGCCCCGCTGCTCAG 1416
QY 1321 GGAGGACCTCAGGTTCTTGAGATTCTCTGACGTGAACCTCATGCCCCGCTCCCAAG 1380
DB 1417 GGAGGACCTCAGGTTCTTGAGATTCTCTGACGTGAACCTCATGCCCCGCTCCCAAG 1476
QY 1381 GACCTGACGTAGAGGGCCCTTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGTC 1440
DB 1477 GACCTGACGTAGAGGGCCCTTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGTC 1536
QY 1441 CGTGCATATCCAGGAGGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTTGGAG 1500
DB 1537 CGTGCATATCCAGGAGGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTTGGAG 1596
QY 1501 CCGGTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCAG 1560
DB 1597 CCGGTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCAG 1656
QY 1561 GCTAGGAGCAAGACAGCCGCTGTCTCCACCTCAGGGCCTTGCTCTGGGCTCCACTTG 1620
DB 1657 GCTAGGAGCAAGACAGCCGCTGTCTCCACCTCAGGGCCTTGCTCTGGGCTCCACTTG 1716
QY 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
DB 1717 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAGCATCTGGGCCAGGGCTCATGGCTGAT 1776
QY 1681 AATTCCCTAGGCTTAACAACCCCAAGCAAGCTTCCGCTCTGTTTATTTTGGTTAA 1740
DB 1777 AATTCCCTAGGCTTAACAACCCCAAGCAAGCTTCCGCTCTGTTTATTTTGGTTAA 1836
QY 1741 CTTATGAAAAATGTATTAAAGAAAGAGTGACGCTGAGAGAGATTCAAGATGGAACACACC 1800
DB 1837 CTTATGAAAAATGTATTAAAGAAAGAGTGACGCTGAGAGAGATTCAAGATGGAACACACC 1896

QY	1801	AGACCCGAGATCACAAGCCAACCATGCCCCAGCCCCCTCCAGCACCCCCCAGCCCGAC	1860
Db	1897	AGACCCGAGATCACAAGCCAACCATGCCCCAGCCCCCTCCAGCACCCCCCAGCCCGAC	1956
QY	1861	CATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTAAACTCATGAAG	1920
Db	1957	CATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTCAACTCATGAAG	2016
QY	1921	ATAACTACCTTCACGTTTGAATAAATATGTTCCCTGTTGAATG	1964
Db	2017	ATAACCACTTCATGTTTGAATAAATATGTTCCCTGTTGAATG	2060

RESULT 3
US-08-253-155A-9

```

1 Patent No. 5691147
2 GENERAL INFORMATION:
3 APPLICANT: Gyuris, Jeno
4 APPLICANT: Draetta, Giulio
5 TITLE OF INVENTION: CDK4 Binding Proteins
6 NUMBER OF SEQUENCES: 95
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: LAHIVE & COCKFIELD
9 STREET: 60 State Street
10 CITY: Boston
11 STATE: MA
12 COUNTRY: USA
13 ZIP: 02109
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: Floppy disk
16 COMPUTER: IBM PC compatible
17 OPERATING SYSTEM: PC-DOS/MS-DOS
18 SOFTWARE: ASCII(text)
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/253,155A
21 FILING DATE: 02-JUN-1994
22 CLASSIFICATION: 435
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Vincent, Matthew P.
25 REGISTRATION NUMBER: 36,709
26 REFERENCE/DOCKET NUMBER: MII-028
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (617) 227-7400
29 TELEFAX: (617) 227-5941
30 INFORMATION FOR SEQ ID NO: 9:
31 SEQUENCE CHARACTERISTICS:
32 LENGTH: 8201 base pairs
33 TYPE: nucleic acid
34 STRANDEDNESS: single
35 TOPOLOGY: linear
36 MOLECULE TYPE: cDNA
37
38 US-08-253-155A-9

```

Query Match	50.9%	Score 999;	DB 2;	Length 8201;
Best Local Similarity	77.8%	Pred. No. 1.2e-263;		
Matches 1332; Conservative	0;	Mismatches 165;	Indels 214;	Gaps 2;

[illegible][illegible][illegible]

Qy	481	GATCGATACGGAACCAAGCAGCGGGAAC	CTACTCCACATCCTCCTGGCATATGAGGAGTAT	540
Db	2174	GATCGATATGGAGCCAAAGCAGAGGGA	CTATTCTACATCCTCCTGGCCATTTCGAGTAT	2233

Oy 541 AACCCGAGGTGGCTACTGCAGGACCTGAGCCACATCGCCGCTTGTTCCTCTCTAT 600
 Db 2234 AACCCGAGGTGGCTACTGCAGGACCTGAGCCACATCACCGCCTTGTTCCTCTTAT 2293

Qy	601	CTTCCTGAGGAGATGCATTCTGGGCACTGTGTGACGCTGCTGGCCAGTGAAGGCACTCC	660
Db	2294	CTGCCTGAGGAGACGCATTCTGGGCACTGTGTGACGCTGCTGGCCAGTGAAGGCACTCC	2353

QY	661	CTGCAGGGATTTCACAGCCCAATGGCGGSA	CGTCCAGGGGCTCCAAGACCAACAGGAG	720
Db	2354	CTGCCAGGATTCCACAGCCCAATGCTGGSA	CAGTCCAGGGGCTCCAAGACCAACAGGAG	2413

```

OY      721 CATGTGTTAGCCACGTCACAACCCAGACCATGGGGCATCAGGACCAAGAAAGATCTATGT 780
      |||||
DB      2414 CATGTGTTAGCCACGTCACAACCCAGACCATGTGGCATCAGGACCAAGGAAGTCTATGC 2473

```

QY 781 GGGCAGTGTCCCGCTTAGCGTCCTCATCCGATATTGATTGACGGGATCTCTCGGG 840

DB 2474 GGGCAGTGTCCCGCTTAGCGTCCTCATCCGGAACCTGATTGACGGGATCTCTCGGG 2533

QY 841 CTCACCCCTGCGCTGTGGACGTATCTGTGTAGAAGCGCAACAGCGCTTGATGCCGATA 900
DB 2534 CTCACCCCTGCGCTGTGGACGTATTTGTGTGAAGAGAACAGCTGTGATGCCAATA 2593

QY 901 ACAGAAATCGCCTTTAAGGTCAGCAGAAAGCGCTCACGAAGACGTCCAGGTGTGGCCCG 960

DB 2594 ACCAGCATGTCTTAAGGTCAGCAGAAAGCGCTCATGAAGACATCCAGGTGTGGCCTG 2653

QY 961 TGGGCACGTTTTCACACCGGTTCCGTTGATACCTGGGCCAGGATGAGSACACTGTGCTC 1020

DB 2654 TGGGCACGTTGGGGAACCAATTCTTCGATACCTGGGCCATGAACGATGACACCGTGCTC 2713

```
QY      1021 AAGCATCTTAGGGCCTCTATGAGAACAATAAAGAAGCAGGGGACTGCCACCCC-- 1078
        |||||
Db      2714 AAGCATCTTAGGGCCTCTACGAAGAATACTAACAGAAGCAAGGGGACTTGCACCCCCA 2773
```

QY	1079	-----	1078
Db	2774	GGCCCAACGCCCTGGGACGAAGTGTGTGSCAGGAAGCCCCCAGCAGTGTGAACCTG	2833

Qy	1079	-----	1078
Db	2834	GGGCGAGTCCGAGGAGCCACCCACCATGCCCCCAACGGCTTCCCATGCCAGGACACAC	2893

Qy	1079	-----	1078
Db	2894	ACCCCTCCCTGGGATCAGCAGACTACAGCGTGTGTCAGTGTAGACCAAGGGGCC	2953

```

QY      1079  -----CAGCCAAACCCGAGCAAGGTCGTGGCATC 1109
          |||||
DB      2954  ACACAGAGACCCCAAGACTCCAGAGATCGAGCCAAACGCGAGCAGGGTCTTGGCACC 3013
          |||||

```

```
QY 1110 CAGGCTGTGCGGCTTACAGTGGCGGAAGACCCTTGCAAGGGGACAGGAGCCCC 1169
    |||
DB 3014 CAGGCTGTGCGGCTTACAGTGGCGGAAGACCCTTGCAAGGGGATATAGCAGGCCCC 3073
QY 1170 TCCAGGCCACAGCCCGGTTCCCGGCCCATTTGGTCAAGTCCCGCCAGGGGACC 1229
    |||
DB 3074 TCCAGGCCACAGCCCGGTTCCAGCGGCCCATTTGCTCAGCTCCCGCCATGGGATC 3133
QY 1230 TCGTTCTTCAACCCCTGTCTGTGGGGCTGTCCGGGAAGACACTACCCTGTGGGAC 1289
    |||
DB 3134 TCGTTTTCACGCCCTGTCTGTGGGGCTGTCCGGGAAGACACTACCCTGTGGGAC 3193
QY 1290 TCAGGTTGCCCCAGCCCGGCTGCTCAAGGAGAACCTCAGGGTTCTTGAGATTCT 1349
    |||
DB 3194 TCAGGTTGCCCCAGCCCTGCGCTCAGGGAAGACCTCAGGGTTCTTGAGATTCT 3253
QY 1350 GCAGTGAACCTCCATGCCCCGCTCCCAACGACCTGACGTAGAGGGCCCTTGTTCCG 1409
    |||
DB 3254 GGAGTGAAGTCAATGCCCGGCTCCCAACGACCTGATATAGGGGGCCCTTGTTCC 3313
QY 1410 CCATTATGATTTACAGACAGAGCTGTGGGTCGTCCTCATATCCAGAGAGACCACTGGC 1469
    |||
DB 3314 CCATTATGATTTGAACGAGAGCTGTGGGTCGTCCTCATATCCAGAGAGACCACTGGC 3373
QY 1470 CCCCTGTGCGAGGCTGAACACCCCTGCCGAG 1500
    |||
DB 3374 CACCTGCTGGCAGGCTGAACACTGCCGAGAG 3404
```

RESULT 4
US-09-513-999C-24808
Sequence 24808, Application US/09513999C
Patent No. 6783961

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 24808
; LENGTH: 121
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-513-999C-24808
```

Query Match 6.2%; Score 121; DB 3; Length 121;
Best Local Similarity 100.0%; Pred. No. 2.6e-23;
Matches 121; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1737 TAAACTTATGAAAATGTATTAAAGAAAGAGTCAGCTCGAGAGAGAGATTCAAGATGGAACA 1796
    |||
DB 1 TAAACTTATGAAAATGTATTAAAGAAAGAGTCAGCTCGAGAGAGAGATTCAAGATGGAACA 60
QY 1797 CACCAAGCCCCAGATCACAAAGCCAAACCATGCCCCAGCCCCCTCCAGACACCCCA 1856
    |||
DB 61 CACCAAGCCCCAGATCACAAAGCCAAACCATGCCCCAGCCCCCTCCAGACACCCCA 120
QY 1857 C 1857
    |
DB 121 C 121
```

RESULT 5
US-09-949-016-2139
Sequence 2139, Application US/09949016
Patent No. 6812339

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2139
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2139
```

Query Match 4.7%; Score 92; DB 3; Length 1925;
Best Local Similarity 60.3%; Pred. No. 8.1e-15;
Matches 152; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
QY 422 ACATCCAGCGCATGACCGGAGCTAAGCGGAGACATTAAAGACATATATTCTCAGGG 481
    |||
DB 118 ACATCAGACAATAGACCTGTGATGCAACCGACATTCGGGACACATATATTAGAG 177
QY 482 ATGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGCGATATGAGAGTATA 541
    |||
DB 178 ACAGATATGTGTTAAGCAACAATCCTTATTCATGTGCTGCTGCTATTTATTA 237
QY 542 ACCCGAGGTGGGCTACTGCGAGGACCTGAGCCACATGCCCCCTGTTCTCTCTATC 601
    |||
DB 238 ACACGGAAGTCGGTATGTGAGGGGATGAGCCAGATCAAGCTTACTCTCATGTATA 297
QY 602 TTCCTGAGGAGATGATTCATTCTGGGCACTGTGCGAGCTGCTGCCAGTGAGAGCACTCCC 661
    |||
DB 298 TGAACGAGGAAGATGCTTCTGGGCCCTGTGTAACCTCTTCTCAGGCCCTTAACATGCCA 357
QY 662 TGCAGGATTTC 673
    |||
DB 358 TGCATGAACATC 369
```

RESULT 6
US-09-513-999C-13267
Sequence 13267, Application US/09513999C
Patent No. 6783961

```
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13267
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
```

```

; LOCATION: 132
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 136
; OTHER INFORMATION: y=c or t
US-09-513-999C-13267

```

```

Query Match      3.8%; Score 74.2; DB 3; Length 187;
Best Local Similarity 97.3%; Pred. No. 2.2e-10;
Matches 73; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 ATGACGTGTAGAGTCGCGGACGTTGGTGCCACAGAGGAGAGACATCATATG 60
      |||
Db      113 ATGACGTGTGTAGAGTCGCGGAGTGTGGTGCCACAGAGGAGAGACATCATATG 172

```

```

QY      61 AAATACGAAAAGGGA 75
      |||
Db      173 AAATACGAAAAGGGA 187

```

RESULT 7

```

US-08-306-691B-14/c
; Sequence 14, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skoraki, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306,691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: NO. 5734039e
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6453 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-306-691B-14

```

```

Query Match      3.8%; Score 74; DB 2; Length 6453;
Best Local Similarity 58.9%; Pred. No. 1.2e-09;
Matches 166; Conservative 0; Mismatches 110; Indels 6; Gaps 2;

```

```

QY      1103 CGGATCCAGGCTGTGCGGCTTCAAGTGGCGGAGAGACCTCTGCAAGGGGAGACAGC 1162
      |||
Db      6200 CGGACCCAGGCTGTGCGGCTTCAAGTGGCGGAGAGACCTCTGCAAGGGGAGAGC 6141

```

```

QY      1163 AGGCCCCCTCAGGCCCCACACGCCGCTTCCCGCGGCCATTGGTACGTTCCCGCCAC 1222
      |||
Db      6140 TGGCTCTCCCTGCCCCATCAGACAGCTCCAGAAATCCCAAGGGTACGT---CCACAT 6084
QY      1223 GGGACCTCGTTCTTCCACACCTGTCTGTGGGCTGTCCGGAGACACCTACCTG 1282
      |||
Db      6083 GGGGCTGTCTCTGCTGACACACTGGCTTAGCTGGGCTGTCCCGAGG---CCAAGGCT 6027
QY      1283 TGGGCACTCAGGCTGTGCCACCGCCGCGCTGCTCAGGAGAGACCTCAGGTTCTGTA 1342
      |||
Db      6026 AGCGGCAATGAGGGGCTTCCGTGAGAGCCAGGGCAGTGGGCTGCTCTGTGAGTCCCTAG 5967
QY      1343 GATTCTGCACTGGAATCCATGCCCCGCTCCCAAGGACC 1384
      |||
Db      5966 GACTCGGGGTGTGCTGCTGCTGCTCCAGATCAGACACACC 5925

```

RESULT 8

```

US-09-209-668-10/c
; Sequence 10, Application US/09209668A
; Patent No. 6114517
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; FILE REFERENCE: ISPH-0336
; CURRENT APPLICATION NUMBER: US/09/209,668A
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1664)..(1774)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2042)..(2220)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2374)..(2533)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3231)..(3350)
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: V00574/Genbank
; DATABASE ENTRY DATE: 1991-01-03
US-09-209-668-10

```

```

Query Match      3.8%; Score 74; DB 3; Length 6453;
Best Local Similarity 58.9%; Pred. No. 1.2e-09;
Matches 166; Conservative 0; Mismatches 110; Indels 6; Gaps 2;

```

```

QY      1103 CGGATCCAGGCTGTGCGGCTTCAAGTGGCGGAGAGACCTCTGCAAGGGGAGACAGC 1162
      |||
Db      6200 CGGACCCAGGCTGTGCGGCTTCAAGTGGCGGAGAGACCTCTGCAAGGGGAGAGC 6141

```

```

QY      1163 AGGCCCCCTCAGGCCCCACACGCCGCTTCCCGCGGCCATTGGTACGTTCCCGCCAC 1222
      |||
Db      6140 TGGCTCTCCCTGCCCCATCAGACAGCTCCAGAAATCCCAAGGGTACGT---CCACAT 6084

```

```

QY      1223 GGGACCTCGTTCTTCCACACCTGTCTGTGGGCTGTCCGGAGAGACCTTACCTTG 1282
      |||
Db      6083 GGGGCTGTCTCTGCTGACACACTGGCTTAGCTGGGCTGTCCCGAGG---CCAAGGCT 6027

```

```

QY      1283 TGGCACTCAGGCTGTGCCACCGCCGCGCTGCTCAGGAGAGACCTCAGGTTCTGTA 1342
      |||
Db      6026 AGCGGCAATGAGGGGCTTCCGTGAGAGCCAGGGCAGTGGGCTGCTCTGTGAGTCCCTAG 5967

```


QY 1343 GATTCCTGACGTGGAAGTCCATGCCCCGCTCCCAACGAGCC 1384
Db 5966 GACTCGGGGTGTGCTGCGCTGTCTCCCAATCAGACACACC 5925

RESULT 9

US-09-356-952-8/c
; Sequence 8, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Boriack-Sjodin, Ann
; APPLICANT: Margarit, S. M.
; APPLICANT: Bor-Sogi, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; EARLIER FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6453
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-356-952-8

Query Match 3.8%; Score 74; DB 3; Length 6453;
Best Local Similarity 58.9%; Pred. No. 1.2e-09;
Matches 166; Conservative 0; Mismatches 110; Indels 6; Gaps 2;

QY 1103 CGGCATCCAGGCTGTGCGGCTTCACTGCGGGAAGACCTCTGCAAGGGGACAGGC 1162
Db 6200 CGGCACCCAGGCTGTGCGGCTTCACTGCGGGAAGACCTCTGCAAGGGGACAGGC 6141
QY 1163 AGGCCCCCTCAGGCCCCCAGGCCCCGCTTCCCGCCATTGTGCTCAAGCTTCCCGCCAC 1222
Db 6140 TGGCTCTCCCTGCTCCATCAGCAGACAGCTCAGATCCCAAGGGGTCAAGT--CCCAT 6084
QY 1223 GGGCAGCTCGTCTTCCACACCCCTGTCTGTGGGCTGTCCGGGAAGACACCTACCTG 1282
Db 6083 GGGGCGCTGTCTGCTCCACACACTGCTGCTGGCTGTCTCCCGAGG--CCAGGCT 6027
QY 1283 TGGGCACTCAGGGGTGTGCGGAGCCCCGCTGCTCAGGGAGGACCTCAGGGTCTTGA 1342
Db 6026 AGCGGCATGAGGGGGCTTCCGTGAGGCCAGGGCAGTGGGCTGTCTCTGAGTCCCTAG 5967
QY 1343 GATTCCTGACGTGGAAGTCCATGCCCCGCTCCCAACGAGCC 1384
Db 5966 GACTCGGGGTGTGCTGCGCTGTCTCCCAATCAGACACACC 5925

RESULT 10

US-09-620-312D-715
; Sequence 715, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing

; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784C1P2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO 715
; LENGTH: 1981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(1585)
US-09-620-312D-715

Query Match 3.7%; Score 72.6; DB 3; Length 1981;
Best Local Similarity 50.5%; Pred. No. 1.7e-09;
Matches 233; Conservative 0; Mismatches 219; Indels 9; Gaps 2;

QY 184 GCGCGGAGGCGGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGATATG 243
Db 266 GCGCTGAGGAAGTACCCCTGAGGTGTGAGGAGAGGAGTCAAGTGGCTGACATG 325
QY 244 CTGGAGACTGGAGAAAT-----ACAAAGCAGCAGAAAGCTATAGATCAGCGTAC 297
Db 326 CTCAACAATCGGAGCAAAATGATGCGCCAAAGACCAAAAGATTCGTGCGGTGCCAA 385
QY 298 AAGGAATGCCATGACATCCGGGCGCCGATGTGTCAAGTCTCTGAACATGAGGA 357
Db 386 AAGGCATCCGCTTCTCTGCGGGCGCTGCTGGCAGTACCTGACGAGCAGGTG 445
QY 358 ATGAAGTGAATAAACCAGGAATACAGATCATGAAGAGAAAGGCAAGAGTCTATCT 417
Db 446 AAGTTACAGCAAAACCTTGAAAGTTGA--CGAGCTGACATGTCCCTGGGAGCC 502
QY 418 GAGCACATCAGCGCATCGACCGGAGCTTAAGCGGACATTAAGAACATATATCTTC 477
Db 503 AAGTGGCTGACGTGATTGAGCTGACCTGACCGGAGTTCATTCATGAATGTTT 562
QY 478 AGGATCGATACGGAACCAAGCAGCGGAATCTACTCCATCTCTGACATGAGAG 537
Db 563 GTGTCCCGGGGGCCACGCGCAGACGACCTATTCGTGTGCTGAAGGCTTACAGCTG 622
QY 538 TATAACCGGAGGTGGCTACTGAGGAGCCTGAGCCATGCGCCGCTGTCTCTC 597
Db 623 TACCGGCCGAGAGGAGGCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 682
QY 598 TATCTTCTGAGGAGGATGATCTTGGGCACTGTGCACT 638
Db 683 CATATGCTCTGAGCAAGCCTTCTGTGCTGTACAGAT 723

RESULT 11

US-09-774-528-336
; Sequence 336, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie

```

; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
US-09-774-528-336
```

Query Match 3.6%; Score 70; DB 3; Length 1823;
Best Local Similarity 49.5%; Pred. No. 8.7e-09;
Matches 209; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

```

QY 220 CGAAGAGCAAGTGGTGATATGCTGGAGACTGGAGAAATACAAAAGCAGAG--A 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 322 CGGAGATGAAGTGGTGAGATGACCTCGACTGGAGAAAACCATGTCCCGCGGTAC 381

QY 277 AAGCTCATAGATCGAGCGTACAGGGAATGCCATGAACATCCGGGCGCCGATGTGTCA 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 382 AAGAAGTAAAGATGACGTCCGAAAGCATCCGCTGTGCCCTGCGCGCCGATGTGTG 441

QY 337 GTCCTCTGAACATTGAGGAATGAAGTTGAAAAACCCCGAAGATACCATCATGAAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 442 CCCCTGTGTGTGGGGCCCATGTGTGCCAGAAAGACGCCCTGGCACTATCAGAGCTG 501

QY 397 GAGAAGGGCAAGGTCATCTGAGCACATCCAGCCGATCGACCCGGAAGCTAAGCGGACA 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 502 GCAGAGGCCCCCTGAGAACCCACAGTGAGTGAACCATTTGGCAGGGACCTGCACCGTCAA 561

QY 457 TTAAGGAGCATATATTCTTCAGGGATCGATACGGAACCAAGCAGGGGAATCTACAC 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 562 TTCCCTCTGACGAGATGTTGTGTGCTCAGGGCCACGGGCAAGGAGGCTCTGTGACG 621

QY 517 ATCCTCTGCGATATGAGGATATTAACCCGAGGTGGCTACTGCAAGGGAAGCTGAGCCAC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 622 GTGCTCAAGGCTTACACCTGTATCGACCGAGCAGGGCTACTGCCAGGCCAGGGGCC 681

QY 577 ATGCGCGCTTGTCTCTCTATCTTCTGAGGAGATGATCTGGGCACTGTGTGACG 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 682 GTGGCTGTGTGTGCTCATGCACCTGCCCCCAGAGAGGAGGCTTGTGTGCTGTGACG 741

QY 637 CT 638
    |
DB 742 AT 743
```

RESULT 12
US-10-120-988-336
; Sequence 336, Application US/10120988
; Patent No. 6919193
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6919193e1 Nucleic Acids and
; TITLE OF INVENTION: Polypeptides

```

; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
US-10-120-988-336
```

Query Match 3.6%; Score 70; DB 3; Length 1823;
Best Local Similarity 49.5%; Pred. No. 8.7e-09;
Matches 209; Conservative 0; Mismatches 210; Indels 3; Gaps 1;

```

QY 220 CGAAGAGCAAGTGGTGATATGCTGGAGACTGGAGAAATACAAAAGCAGAG--A 276
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 322 CGGAGATGAAGTGGTGAGATGACCTCGACTGGAGAAAACCATGTCCCGCGGTAC 381

QY 277 AAGCTCATAGATCGAGCGTACAGGGAATGCCATGAACATCCGGGCGCCGATGTGTCA 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 382 AAGAAGTAAAGATGACGTCCGAAAGCATCCGCTGTGCCCTGCGCGCCGATGTGTG 441

QY 337 GTCCTCTGAACATTGAGGAATGAAGTTGAAAAACCCCGAAGATACCATCATGAAG 396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 442 CCCCTGTGTGTGGGGCCCATGTGTGCCAGAAAGACGCCCTGGCACTATCAGAGCTG 501

QY 397 GAGAAGGGCAAGGTCATCTGAGCACATCCAGCCGATCGACCCGGAAGCTAAGCGGACA 456
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 502 GCAGAGGCCCCCTGAGAACCCACAGTGAGTGAACCATTTGGCAGGGACCTGCACCGTCAA 561

QY 457 TTAAGGAGCATATATTCTTCAGGGATCGATACGGAACCAAGCAGGGGAATCTACAC 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 562 TTCCCTCTGACGAGATGTTGTGTGCTCAGGGCCACGGGCAAGGAGGCTCTGTGACG 621

QY 517 ATCCTCTGCGATATGAGGATATTAACCCGAGGTGGCTACTGCAAGGGAAGCTGAGCCAC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 622 GTGCTCAAGGCTTACACCTGTATCGACCGAGCAGGGCTACTGCCAGGCCAGGGGCC 681

QY 577 ATGCGCGCTTGTCTCTCTATCTTCTGAGGAGATGATCTGGGCACTGTGTGACG 636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 682 GTGGCTGTGTGTGCTCATGCACCTGCCCCCAGAGAGGAGGCTTGTGTGCTGTGACG 741

QY 637 CT 638
    |
DB 742 AT 743
```

RESULT 13
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-F18
;
US-08-232-463-14
```

```

Query Match      3.5%; Score 68.4; DB 2; Length 7218;
Best Local Similarity 4.5%; Pred. No. 4.4e-08;
Matches 18; Conservative 231; Mismatches 147; Indels 0; Gaps 0;
```

```

QY      38 AAGACGAGAGCATCATTTATGAATACGAAAGGAGACCGAGCTGGCTCCAGAGG 97
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      1456 AAGAGATAGAGAAATTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1397
QY      98 ACAAGGGGCTAAGCCTTTTGAAGCTACACACAACGTCATGATTGGGATTGTAC 157
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1396 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1337
QY      158 ATGAGACGAGCTGCTCTCTGACTGCGGAGGCGAAGCAATTGCGCGGAGATCA 217
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1336 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1277
QY      218 GCCGAAAGACAGTGGTGATATGCTGGAGACTGGGAGAAATACAAAGCAGAGAA 277
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1276 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1217
QY      278 AGCTCATAGATCGAGCGTACAGGGAATGCCATGAACATCCGGGGCCGATGTGTCAG 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1216 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1157
QY      338 TCCTCTGAACATTGAGGAATGAAGTGAAGAAACCCCGAAGATACCATGATGAGAG 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1156 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1097
QY      398 AGAAGGGCAGAGGTGATCTGAGCAGATCCAGCGCA 433
      : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      1096 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1061
```

```

RESULT 14
US-09-270-767-9899/c
; Sequence 9899, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9899
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
;
US-09-270-767-9899
```

```

Query Match      3.2%; Score 63.4; DB 3; Length 806;
Best Local Similarity 55.1%; Pred. No. 3.9e-07;
Matches 124; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
```

```

QY      433 ATGACCGGAGCGTAAGCGGACATTAAGAGCATATATTCTCAGGATCGATACGA 492
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      668 ATCCAGCGGACATCCATCGACTTCCCGGCGCACAAATGCTTCAAGAAATTGGCGT 609
QY      493 ACCAAGCAGCGGAACTACTCCACATCTCTCTGCAATGAGAGTAAACCCGAGGTG 552
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      608 TCTGGCCAGATGCACCTTTTAAAGTGTCCAAAGCGTATGCGCTTCATGACAGCGAGTT 549
QY      553 GGCTACTGAGGAGACCTGAGCCATCGCCGCTTGTCTCTCTATCTTCTGAGAG 612
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      548 GGATATTGTCAGGGTCTAAGTTTCATAGCAGCTAGTCTGCTCTTCATATGCTGAGAG 489
QY      613 GATGATTTGGGCACTGTGTGCACTGCTGCGCAGTGAGAGCAC 657
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      488 GATGCGTTGTGTCTGTGTAGCGCTTATGTACGACTATGGGCTC 444
```

RESULT 15

```

US-08-363-300-1
; Sequence 1, Application US/08363300
; Patent No. 5700927
```

GENERAL INFORMATION:

```

; APPLICANT: Zon, Leonard and Richardson, Paul
; TITLE OF INVENTION: Tbc1 Gene and Uses Thereof
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/363,300
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 04590/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4039 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..3829
;
US-08-363-300-1
```

```

Query Match      2.7%; Score 53.6; DB 2; Length 4039;
Best Local Similarity 49.6%; Pred. No. 0.00039;
```

		Matches	137;	Conservative	0;	Mismatches	139;	Indels	0;	Gaps	0;
QY	367	AAAAACCCCGGAAGATACCAGATCATGAAGAGGCAAGGTCATCTGAGCACATC	426								
Db	2752	AAACAGCAGCCAAAGAGACGTGCCCTACAAAGAGCTCCTGAAGAGCTGACCTCGCAGCAG	2811								
QY	427	CAGCGCATCGACCGGAGCGTAAGCGGACATTAAGAGACATATATCTTCAGGGATCGA	486								
Db	2812	CAGCCATCTCTCATCGACCTCGGCGAACCTTCCAACACATCCATCTCTCTGCCCGAG	2871								
QY	487	TACGGAACCAAGCAGCGGGAATCTCCACATCCTCTGCAATATGAGAGTATAACCG	546								
Db	2872	CTTGAGCAGGTCTAGCTGTCACTTACACATTTCTGAAGGCCCTACTCGCTTCTGACCGAG	2931								
QY	547	GAGGTGGCTTACTTGCAGGGACCTGAGCCACATCGCCGCTTGTCTCTCTATCTTCT	606								
Db	2932	GAGGTGGATACTGCAAGGTCTCAGCTTGTGTGGCAGGCATTTTGCTTCTTCACATGAGT	2991								
QY	607	GAGGAGATGCATTTCTGGGCACTGTGTGCACTGCTG	642								
Db	2992	GAGGAAGAGGCGTTCAAGATGCTCAAGTTCCTGATG	3027								

Search completed: April 5, 2006, 14:01:08
Job time : 365 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:32:54 ; Search time 5318 Seconds
(without alignments)
17279.020 Million cell updates/sec

Title: US-10-071-838-1
Perfect score: 1964
Sequence: 1 atggacgtggtagaggtcgc.....aatgttctctgtgaatg 1964

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_hic:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_est7:*

9: gb_gse1:*

10: gb_gse2:*

11: gb_gse3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1948	99.2	2072	4	HSM801828
2	1585.8	80.7	2497	4	BC047739
3	859.8	43.8	986	5	BX346181
4	841.2	42.8	915	5	BQ923700
5	814.6	41.5	909	3	BI753688
6	797.6	40.6	954	5	BU526609
7	797.4	40.6	919	3	BI522600
8	783.6	39.9	868	6	CD653466
9	751.6	38.3	881	5	BU542901
10	741.2	37.7	818	3	BI597531
11	732.6	37.3	930	5	BQ920917
12	730.2	37.2	1013	3	BM560320
13	724.8	36.9	839	3	BI601698
14	718.2	36.6	1141	3	BM553146
15	713.6	36.3	831	3	BI828272
16	707.4	36.0	709	2	BI088323
17	701.8	35.7	1079	3	BM563571
18	700.8	35.7	732	6	CA748642
19	696.6	35.5	727	1	AL042933
20	694.4	35.4	946	5	BQ891586
21	686.6	35.0	875	9	AQ788033
22	675	34.4	741	5	BU688833

ALIGNMENTS

23	663.8	33.8	667	3	BI827136	BI827136	603077337
24	651.4	33.2	839	6	CA487310	CA487310	AGENCOURT
25	645.8	32.9	840	3	BI829764	BI829764	603079835
26	641	32.6	687	6	CB241156	CB241156	UI-CF-FN0
27	640.4	32.6	882	3	BI520522	BI520522	603071531
28	635.2	32.3	680	2	BE676151	BE676151	7F24804.x
29	632.4	32.2	634	7	CN262008	CN262008	170005977
30	632	32.2	964	3	BM914524	BM914524	AGENCOURT
31	628.4	32.0	630	3	BM842553	BM842553	K-EST0119
32	627.4	31.9	629	3	BM843014	BM843014	K-EST0120
33	626.8	31.9	669	1	AI867050	AI867050	tz45e05.x
34	626.4	31.9	628	3	BM762465	BM762465	K-EST0043
35	625.4	31.8	654	5	BU625940	BU625940	UI-H-FG1-
36	622.4	31.7	742	3	BI832051	BI832051	603075809
37	620	31.6	628	8	DN999748	DN999748	TC104423
38	610.2	31.1	633	8	DR006725	DR006725	TC105000
39	609.6	31.0	700	1	AQ420821	AQ420821	RPCI-11-2
40	609.4	31.0	700	1	AI559979	AI559979	tg77f06.x
41	604.4	30.8	655	5	BU607318	BU607318	UI-CF-FN0
42	603.4	30.7	766	3	BQ007813	BQ007813	UI-H-EI0-
43	600	30.5	637	1	AM440435	AM440435	hb90e11.x
44	595	30.2	802	7	CR766757	CR766757	DKFZp468D
45	591.8	30.1	650	3	BI830812	BI830812	603080926

RESULT 1
HSM801828
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp434P2235 (from clone DKFZp434P2235).
ACCESSION AL136860
VERSION AL136860.1 GI:12053224
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2072)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German CDNA Consortium
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project. This clone (DKFZp434P2235) is available at
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in
Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp434P2235
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/
location/Qualifiers
1. 2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434P2235"
/db_xref="taxon:9606"
/clone="DKFZp434P2235"
/tissue_type="testis"
/clone_lib="434 (synonym: hte3j). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="TBC1 domain family, member 3"
1. 2072
/gene="DKFZp434P2235"
99. 1748
CDS

/gene="DKF2p434P2235"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66794.1"
/db_xref="GI:12053225"
/db_xref="UniProt/Swiss-Prot:Q81ZP1"
translation="MDVVEVAGSWMAQERDIIIMKYEKHRAGLPEDKGPFRSYNN
NVDHLGIVHETELPPLTAREAKQIRREISRSKSWMDMLDWEKYKSRKLDIRAYKGM
PMNIRGPMWSVLNTEEMKLNDRGQIIMEKSKSEHIQRIIDRVSGTIRKHIIFR
DRYTRQRELIHLILAYEENPEVGYCRDLSHIALFLIYLPEDAFWALVQLASER
HSLQGFHSPNGSTVQGLQDQEHVATSQPKTMGHQDKDLGQCSPLGCLIRLLDG
ISLGLTLRLMDVYLVEGEQALMPTRIAFKVQOKRLTTSRCGPWARFCNRFVDTWAR
DEDTVLKLRLASMKLTKRKGDLPPPAKPEQSSASRPVPASRGKTLCKGRQAPPG
PPARFPRPIWSASPPRAPRSSTPCPGAVREDTYPVGTQGVSPALAOGGQGSWRL
QWNSMPRLPTDLDBEGPWFRRHYDFRQSCWRAISQEDQLAPCWQAEHPAERYRSAPAA
PSTDSDQGTFFRARDQQAFTSGPCLGHLLESQFPPEGF"

ORIGIN

Query Match 99.2%; Score 1948; DB 4; Length 2072;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGACGTGTAGAGGTGCGCGGACGTTGTGGGCACAAGACGAGACATCATTAATG 60
DB 99 ATGACGTGTAGAGGTGCGCGGACGTTGTGGGCACAAGACGAGACATCATTAATG 158
QY 61 AAATACGAAAGGACACCGAGCTGGGCTGCAGAGACAAAGGGCCTAAGCCTTTTGA 120
DB 159 AAATACGAAAGGACACCGAGCTGGGCTGCAGAGACAAAGGGCCTAAGCCTTTTGA 218
QY 121 AGCTACAACAACGTCGATCATTTGGGGATTGTACATGACGAGCTGCCTCCTG 180
DB 219 AGCTACAACAACGTCGATCATTTGGGGATTGTACATGACGAGCTGCCTCCTG 278
QY 181 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 240
DB 279 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 338
QY 241 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAAAAGCTCATAGATCGAGCGTACAG 300
DB 339 ATGCTGGAGACTGGGAGAAATACAAAAGCAGAAAAGCTCATAGATCGAGCGTACAG 398
QY 301 GGAATGCCCATGAATCCGGGGCCCGATGTGTAGTCTCTCTGAACATTGAGAAATG 360
DB 399 GGAATGCCCATGAATCCGGGGCCCGATGTGTAGTCTCTCTGAACATTGAGAAATG 458
QY 361 AAGTTGAAAAACCCCGGAAGATACCAAGATGAGAAGGCAAGAGTCACTGAG 420
DB 459 AAGTTGAAAAACCCCGGAAGATACCAAGATGAGAAGGCAAGAGTCACTGAG 518
QY 421 CACATCCAGCGCATCGACCGGGACGTAAAGGGGACATTAAAGAACATATATTCTTCAAG 480
DB 519 CACATCCAGCGCATCGACCGGGACGTAAAGGGGACATTAAAGAACATATATTCTTCAAG 578
QY 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 540
DB 579 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAC 638
QY 541 AACCCGAGGTGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
DB 639 AACCCGAGGTGGCTACTGCAAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 698
QY 601 CTTCTGAGGAGATGCATTCTGGGACATGTGCAAGTGTGCGCAAGTGAAGGACATCC 660
DB 699 CTTCTGAGGAGATGCATTCTGGGACATGTGCAAGTGTGCGCAAGTGAAGGACATCC 758
QY 661 CTGCAAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
DB 759 CTGCAAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 818
QY 721 CATGTGGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 780
DB 819 CATGTGGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 878

QY 781 GGGCAGTGTCCCGCTTAGGCTGCCTCATCCGGATATGATTGACGGGATCTCTCGGG 840
DB 879 GGGCAGTGTCCCGCTTAGGCTGCCTCATCCGGATATGATTGACGGGATCTCTCGGG 938
QY 841 CTCACCCCTGCGCCTGTGGGACGTGTATCTGTGAGAAGCGCAACAGCGCTTGATGCCGATA 900
DB 939 CTCACCCCTGCGCCTGTGGGACGTGTATCTGTGAGAAGCGCAACAGCGCTTGATGCCGATA 998
QY 901 ACAAGATCGCCTTTAAGGTTACAGAGAAGCGCCTCAGAGAAGCGTCCAGGTGTGGCCCG 960
DB 999 ACAAGATCGCCTTTAAGGTTACAGAGAAGCGCCTCAGAGAAGCGTCCAGGTGTGGCCCG 1058
QY 961 TGGGACGTTTTTGCAAACCGGTTGTTGATACCTGGGCCAGGATGAGACATGTGCTC 1020
DB 1059 TGGGACGTTTTTGCAAACCGGTTGTTGATACCTGGGCCAGGATGAGACATGTGCTC 1118
QY 1021 AAGCATCTTAGGGCCTCTATGAAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 1080
DB 1119 AAGCATCTTAGGGCCTCTATGAAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCA 1178
QY 1081 GCCAAACCCGACGAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAAG 1140
DB 1179 GCCAAACCCGACGAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAAG 1238
QY 1141 ACCCTTGAAGGGGGACAGGCAAGCCCTCCAGGCCCAACAGCCCGTTCCCGCGGCC 1200
DB 1239 ACCCTTGAAGGGGGACAGGCAAGCCCTCCAGGCCCAACAGCCCGTTCCCGCGGCC 1298
QY 1201 ATTTGTGACGTTCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT 1260
DB 1299 ATTTGTGACGTTCCCGCCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT 1358
QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320
DB 1359 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1418
QY 1321 GGAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCCATGCCCGGCTTCCCAAG 1380
DB 1419 GGAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCCATGCCCGGCTTCCCAAG 1478
QY 1381 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTACAGACAGAGCTGTGGGTC 1440
DB 1479 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTTACAGACAGAGCTGTGGGTC 1538
QY 1441 CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGCGGACAGGCTGAACACCCCTCGGAG 1500
DB 1539 CGTGCCATATCCAGAGGACCAAGCTGGCCCCCTGTGCGGACAGGCTGAACACCCCTCGGAG 1598
QY 1501 CGGTGAGATCGGCTTTCGTGCAACCCAGCACTGATTCGAACAGGACACCCCTTCAGA 1560
DB 1599 CGGTGAGATCGGCTTTCGTGCAACCCAGCACTGATTCGAACAGGACACCCCTTCAGA 1658
QY 1561 GCTAGGACGAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTGCGGCTTCACTTG 1620
DB 1659 GCTAGGACGAACAGCAAGTGTGCTCCACCTCAGGGCCTTGCTGCGGCTTCACTTG 1718
QY 1621 GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATTTGGGCCAGGCTCATGGCTGAT 1680
DB 1719 GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATTTGGGCCAGGCTCATGGCTGAT 1778
QY 1681 AATTTCCCTAAGGCTTAACAACCAAGCAAGCTTCCGCTCTCGTTTATTTTGTGTTAA 1740
DB 1779 AATTTCCCTAAGGCTTAACAACCAAGCAAGCTTCCGCTCTCGTTTATTTTGTGTTAA 1838
QY 1741 CTTATGAATAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
DB 1839 CTTATGAATAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1898
QY 1801 AGACCCAGATCACAAAGCCAAAGCATGCCCAGGCCCTCCAGCACACCCCAAGCCACGAC 1860
DB 1899 AGACCCAGATCACAAAGCCAAAGCATGCCCAGGCCCTCCAGCACACCCCAAGCCACGAC 1958

QY 1861 CATGTTCTGAATTCGTACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
|||||
DB 1959 CATGTTCTGAATTCGTACGACACCGTGAGCCTGCTTGTACTTTCAAACTCATGGAAG 2018

QY 1921 ATAACACTTCACGTTTGAATAAATGTTCTCTGTTGAATG 1964
|||||
DB 2019 ATAACCACTTCATGTTTGAATAAATGTTCTCTGTTGAATG 2062

RESULT 2
BC047739

LOCUS BC047739 2497 bp mRNA linear HTC 20-MAR-2003
DEFINITION Homo sapiens, similar to Rab GTPase-activating protein PRCL7, clone
IMAGE:5743752, mRNA.

ACCESSION BC047739 GI:29126829
VERSION BC047739.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2497)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAK Plate: 98 Row: d Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984
This clone has the following problem: retained intron.

FEATURES
source 1..2497
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5743752"
 /tissue_type="Brain, adult medulla"
 /clone_lib="NIH_MGC_119"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"

ORIGIN
Query Match 80.7%; Score 1585.8; DB 4; Length 2497;
Best Local Similarity 89.8%; Pred. No. 0;
Matches 1796; Conservative 0; Mismatches 12; Indels 193; Gaps 1;

QY 157 CATGACGGAAGCTGCTCTCTGACTGCGCGGAGGCGAACAATTGCGGAGATC 216
|||||
DB 468 CGTAGACGGAAGCTGCTCTCTGACTGCGCGGAGGCGAACAATTGCGGAGATC 527

QY 217 AGCCGAAGACCAAGTGGTGAATATGCTGGAGACTGGGAGAAATACAAAAGCAGACA 276
|||||
DB 528 AGCCGAAGACCAAGTGGTGAATATGCTGGAGACTGGGAGAAATACAAAAGCAGACA 587

QY 277 AAGCTCATATGATCGAGCGTACAAAGGAATGCCCATGAACATCCGGGCGCATGTGTCA 336

DB 588 AAGCTCATATGATCGAGCGTACAAAGGAATGCCCATGAACATCCGGGCGCATGTGTCA 647
|||||
QY 337 GTCCTCTGACATTGAGGAATGAAGTGAACAAACCCCGAAGATACCAGATCATGAAG 396
648 GTCTCTCTGAACATTGAGGAATGAAGTGAACAAACCCCGAAGATACCAGATCATGAAG 707

QY 397 GAGAAAGGCAAGAGTCACTGAGCAGATCCAGCGCATGACCGGACGTAAGCGGACA 456
708 GAGAAAGGCAAGAGTCACTGAGCAGATCCAGCGCATGACCGGACGTAAGCGGACA 767

QY 457 TTAAGGAAGCATATATTCTTCAGGGATCGATACGAAACCAAGCAGCGGAATACTCCAC 516
768 TTAAGGAAGCATATATTCTTCAGGGATCGATACGAAACCAAGCAGCGGAATACTCCAC 827

QY 517 ATCCTCGGCATATGAGAGTATAACCCGAGGTGGGCTTACTGACGGACCTGAGCCAC 576
828 ATCCTCGGCATATGAGAGTATAACCCGAGGTGGGCTTACTGACGGACCTGAGCCAC 887

QY 577 ATGCGCGCTTGTCTCTCTCTATCTTCTTGAGAGAGATGCAATTGCGCACTGTGACG 636
888 ATGCGCGCTTGTCTCTCTCTATCTTCTTGAGAGAGATGCAATTGCGCACTGTGACG 947

QY 637 CTGCTGGCAGTGAAGGCACTCCCTGACGGATTTCACAGCCCAATGCGGACCGTC 696
948 CTGCTGGCAGTGAAGGCACTCCCTGACGGATTTCACAGCCCAATGCGGACCGTC 1007

QY 697 CAGGGCTCCAAAGACCAACAGAGCATGTGTAGCCACGTCAACCCCAAGACCATGGG 756
1008 CAGGGCTCCAAAGACCAACAGAGCATGTGTAGCCACGTCAACCCCAAGACCATGGG 1067

QY 757 CATCAGACAAAGAAATCTATGTGGCAGTGTCCCGCTTAGGCTGCTCATCCGATA 816
1068 CATCAGACAAAGAAATCTATGTGGCAGTGTCCCGCTTAGGCTGCTCATCCGATA 1127

QY 817 TTGATTGACGGGATCTCTCTCGGGCTCACCTGCGCTGTGGGACGTGTATCTGTAGAA 876
1128 TTGATTGACGGGATCTCTCTCGGGCTCACCTGCGCTGTGGGACGTGTATCTGTAGAA 1187

QY 877 GGCGAAGCGGCTGATGCCGATTAACAAGATCGCTTTAAGGTTACAGACA----- 928
1188 GGCGAAGCGGCTGATGCCGATTAACAAGATCGCTTTAAGGTTACAGAGTAAGTCT 1247

QY 929 ----- 928
1248 ACGTGCCCAAGCGGCGCTGGGAGCCCTGGGTCAACCCCGACTGCGCCGAGGGCAG 1307

QY 929 ----- 928
1308 CTTCTTCACACTGTCTCATGATCTCTGTCTGCGCCAGAGGGAGGTCTGGCCAGGTG 1367

QY 929 ----- 928
1368 GCTGGGAGGACACTGTGACACCGAGCCCATCCCCACATGACCCAGATGAAGTGCAGA 1427

QY 929 -----AGCGCCTCAGGAAGACGTTCAGGTGTGGCCCGTGGGCAAGTTTGGCAACGGTT 983
1428 GTGTAGCGCCTCAGGAAGACGTTCAGGTGTGGCCCGTGGGCAAGTTTGGCAACGGTT 1487

QY 984 CGTTGATACCTGGGCCAGGAGTGAAGACACTGTGCTCAAGCATCTTAGGGCTCTATGAA 1043
1488 CGTTGATACCTGGGCCAGGAGTGAAGACACTGTGCTCAAGCATCTTAGGGCTCTATGAA 1547

DB 1548 GAAACTTAACAAGAAAGAGGGGAGCTTGCCACCCCGCAAAACCCGAGCAAGGTCGTC 1607

QY 1104 GGCATCAAGGCTGTGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGGAACAGGCA 1163
1608 GGCATCAAGGCTGTGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGGAACAGGCA 1667

QY 1164 GGCCCTTCCAGGGCCACCAAGCGCGGTTCCCGCGGCCATTGTGTCAGCTTCCCGCCACG 1223
|||||

Db 1668 GGGCCCTCCAGGCCCCACCAGCCCGGTTCCCGGGCCCATTTGGTCAGCTTCCCTGCCAG 1727

Qy 1224 GGCACCTCGTTCTTCCACACCCCTGTCTGGTGGGCTGTCCGGAGAACAACCTACCTGT 1283

Db 1728 GGCACCTCGTTCTTCCACACCCCTGTCTGGTGGGCTGTCCGGAGAACAACCTACCTGT 1787

Qy 1284 GGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAGGAGGACCTCAGGGTTCCTGAG 1343

Db 1788 GGGCACTCAGGGTGTGCCAGCCCGGCTGGCTCAGGAGGACCTCAGGGTTCCTGAG 1847

Qy 1344 ATTCCTCAGTGGAACTCCATGCCCCGCTCCCAACGACCTGACGTAGAGGGCCCTG 1403

Db 1848 ATTCCTCAGTGGAACTCCATGCCCCGCTCCCAACGACCTGACGTAGAGGGCCCTG 1907

Qy 1404 GTTCCGCCATTATGATTTCAAGACAGAGCTGCTGGTCCGTCCATATCCAGAGAGCA 1463

Db 1908 GTTCCGCCATTATGATTTCAAGACAGAGCTGCTGGTCCGTCCATATACAGAGAGCA 1967

Qy 1464 GCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGAGCGGGTGAATCGGCTTCGCTGC 1523

Db 1968 GCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGAGCGGGTGAATCGGCTTCGCTGC 2027

Qy 1524 ACCCAGCACTGATTCGACAGGCGACCCCTTCAGAGCTAGGAGCAACAGCCGCTGC 1583

Db 2028 ACCCAGCACTGATTCGACAGGCGACCCCTTCAGAGCTAGGAGCAACAGCACTGTC 2087

Qy 1584 TCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTGAAAGTCTCAGTTCCCTCCAG 1643

Db 2088 TCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTGAAAGTCTCAGTTCCCTCCAG 2147

Qy 1644 CTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACC 1703

Db 2148 CTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGATATTTCCCTAGGCTTAACAACC 2207

Qy 1704 AAGCAAGCTTCGGCTCTCTGTTTATTTTGGTTAACTTATGAAAATGTATTAAGAA 1763

Db 2208 AAGCAAGCTTCGCATCTCTGTTTATTTTGGTTAACTTATGAAAATGTATTAAGAA 2267

Qy 1764 AGTGCACTCGAGAGAGATTCAAGATGGAACAACACAGACCCAGATCACAAGCCAAC 1823

Db 2268 AGTGCACTCGAGAGAGATTCAAGATGGAACAACAACAGACCCAGATCACAAGCCAAC 2327

Qy 1824 CATGCCAGCCCTCCAGCAACCCCAAGCCCAAGACCATGTTCTGAATTCGACGACA 1883

Db 2328 CATGCCAGCCCTCCAGCAACCCCAAGCCCAAGACCATGTTCTGAATTCGACGACA 2387

Qy 1884 CCGTAGCCCTGCTTGTACTTTAACTCATGGAAGATACTTCAAGTTTGA 1943

Db 2388 CCGTAGCCCTGCTTGTACTTTAACTCATGGAAGATACTTCAAGTTTGA 2447

Qy 1944 TAAATGTTTCTGTGAATG 1964

Db 2448 TAAATGTTTCTGTGAATG 2468

RESULT 3

LOCUS BX346181/c 986 bp mRNA linear EST 08-APR-2004

DEFINITION BX346181 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED

ACCESSION Homo sapiens cDNA clone CS0DJ010YK20 3-PRIME, mRNA sequence.

VERSION BX346181

KEYWORDS BX346181.1 GI:30369074

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE Homo sapiens

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

JOURNAL Homnidae; Homo.

COMMENT 1 (bases 1 to 986)

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8800.r

For more information about this cluster, see <http://www.genoscope.cns.fr/cdna?c=CS0A0J010BF10NP1&c=8800.r>.

FEATURES

location/Qualifiers

1..986

1..986

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DJ010YK20"

/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/cell_line="JURKAT"

/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 43.8%; Score 859.8; DB 5; Length 986;

Best Local Similarity 95.4%; Pred. No. 4.9e-219;

Matches 925; Conservative 9; Mismatches 29; Indels 7; Gaps 5;

Qy 971 TTGCAACGGGTCTGTGATACCTGGGCCAGGATGAGCACTGTGCTCAAGACTTTA 1030

Db 979 KTTTSAACCGTCTCTATATACCTGGCCMAGGATGAGRACACTGT-CTMAAGCATCTTA 921

Qy 1031 GGGCTTATGAAGAACTAACAAGAGAGGGGACTGACCACCCCAAGCCAAACCCG 1090

Db 920 GGCC--TCTAKAABAATAACAAGAAAGCAGGGACC--KCAACCCCAACCAACCCG 865

Qy 1091 AGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAAGACCTCTGCA 1150

Db 864 AGCAAGGTCGTGGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAAGACCTCTGCA 805

Qy 1151 AGGGGACAGGCAAGGCCCTCCAGGCCCAAGAGCCCGGTTCCCGGCCATTGGTCA 1210

Db 804 AGGGGACAGGCAAGGCCCTCCAGGCCCAAGAGCCCGGTTCCCGGCCATTGGTCA 745

Qy 1211 CTTCCCCGCAAGGCACTGCTTCTTCCACACCTGTCTGTGGGCTGTCCGGGAAG 1270

Db 744 CTTCCCCGCAAGGCACTGCTTCTTCCACACCTGTCTGTGGGCTGTCCGGGAAG 685

Qy 1271 ACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCCCTGGCTCAGGAGGACCTC 1330

Db 684 ACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCCCTGGCTCAGGAGGACCTC 625

Qy 1331 AGGTTCTGAGATTCTCGACGTGAACTCCATGCCCCGCTCCCAACGAGACTGACG 1390

Db 624 AGGTTCTGAGATTCTCGACGTGAACTCCATGCCCCGCTCCCAACGAGACTGACG 565

Qy 1391 TAGAGGGCCCTGTGCTCCGCCATTATGATTTCAAGACAGAGCTGCTGGTCCGTCATAT 1450

Db 564 TAGAGGGCCCTGTGCTCCGCCATTATGATTTCAAGACAGAGCTGCTGGTCCGTCATAT 505

Qy 1451 CCCAGGAGCAAGCTGGCCCCCTGCTGGCAAGGCTGAACACCTGCGAGCGGGTGAAT 1510

Db 504 CCCAGGAGCAAGCTGGCCCCCTGCTGGCAAGGCTGAACACCTGCGAGCGGGTGAAT 445

Qy 1511 CGGCTTCTGTCACCCAGCACTGATTCGACAGGCGACCCCTTCAGAGCTAGGAGCG 1570

Db 444 CGGCTTCTGTCACCCAGCACTGATTCGACAGGCGACCCCTTCAGAGCTAGGAGCG 385

Qy 1571 AACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTCCACTTGAAAGTTCTC 1630

Db 384 AACAGCCGTGTGCTCCACCTCAGGGGCTTGCTCTGCGGCTCCACTTGAAAGTTCTC 325

QY 1631 AGTCCCTCCAGGCTTCTAGAAGCATTGGGCCAGGCTCATGGCTGATAATTCCCTA 1690

Db 324 AGTCCCTCCAGGCTTCTAGAAGCATTGGGCCAGGCTCATGGCTGATAATTCCCTA 265

QY 1691 GGCTTAACAACCCAAAGCAAGCTTGCCTCTGTTTATTTGGTTAACTTATGAAA 1750

Db 264 GGCTTAACAACCCAAAGCAAG-TTCGCATCTCGTTTATTTGGTTAACTTATGAAA 206

QY 1751 TGTATTAAGAAAGAGTGCAGCTCGAGAGATTTCAGAGATGGAACACACGAGCCGAGA 1810

Db 205 TGTATTAAGAAAGAGTGCAGCTCGAGAGATTTCAGAGATGGAACACACGAGCCGAGA 146

QY 1811 TCACAAAGCCAAACCATGCCCCAGCCCTCCAGACACCCCGACGACCATCGTTCTG 1870

Db 145 TCACAAAGCCAAACCATGCCCCAGCCCTCCAGACACCCCGACGACCATCGTTCTG 86

QY 1871 AATTCTGACGACACCGTGAGCCCTGCTTGTACTTTAACTC-ATGGAAGATACTACC 1929

Db 85 AATTCTGACGACACCGTGAGCCCTGCTTGTACTTTAACTCNAATGGAAGATAACCNMC 26

QY 1930 TTCAACGTTTT 1939

Db 25 CTTCNATGTT 16

RESULT 4

LOCUS BQ923700 915 bp mRNA linear EST 20-AUG-2002

DEFINITION AGENCOURT_8798484 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6425622

5', mRNA sequence.

ACCESSION BQ923700

VERSION BQ923700.1 GI:22338731

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 915)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM2607 row: 1 column: 07
High quality sequence stop: 639.

FEATURES

source

1. 915

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6425622"

/tissue_type="epidermoid carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC 101"

/note="Organ: lung; Vector: pOTB7, Site_1: EcoRI, Site_2: XhoI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 42.8%; Score 841.2; DB 5; Length 915;
Best local Similarity 98.5%; Pred. No. 4.7e-214;
Matches 890; Conservative 0; Mismatches 10; Indels 4; Gaps 4;

QY 409 AGTCATCTGAGACATCCAGCGCATCGACCGGACGTAACGGGACATTAAAGAACAT 468

Db 1 AAGGCATCTGAGACATCCAGCGCATCGACCGGACGTAACGGGACATTAAAGAACAT 60

QY 469 ATATTCTTCAAGGATCGATACGGAACCAAGCAGCGGGAATCTCCACATCTCTGSCA 528

Db 61 ATATTCTTCAAGGATCGATACGGAACCAAGCAGCGGGAATCTCCACATCTCTGSCA 120

QY 529 TATGAGAGTATAACCCGAGGTGGCTACTGCAAGGACCTGACCATCGCCGCTTG 588

Db 121 TATGAGAGTATAACCCGAGGTGGCTACTGCAAGGACCTGACCATCGCCGCTTG 180

QY 589 TTCCTCCTTATCTTCTGAGAGGATGCAATTCTGGGCACTGGTGACGCTGGCGAGT 648

Db 181 TTCCTCCTTATCTTCTGAGAGGATGCAATTCTGGGCACTGGTGACGCTGGCGAGT 240

QY 649 GAGAGCACTCCCTGCAAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGCTCCA 708

Db 241 GAGAGCACTCCCTGCAAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGCTCCA 300

QY 709 GACCAACAGGACATGTGTAGCCAGTCAACAACCCAAAGCATGGGCAATCAGACAAG 768

Db 301 GACCAACAGGACATGTGTAGCCAGTCAACAACCCAAAGCATGGGCAATCAGACAAG 360

QY 769 AAAGATCTATGTGGGAGTGTCCCGTTAGGCTGCTCATCCGATATTGATGACGG 828

Db 361 AAAGATCTATGTGGGAGTGTCCCGTTAGGCTGCTCATCCGATATTGATGACGG 420

QY 829 ATCTCTCGGGCTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAGCGAAGCGG 888

Db 421 ATCTCTCGGGCTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAGCGAAGCGG 480

QY 889 TTGATGCCGATTAACAAGATCGCTTTAAGGTTACGACGAACCGCTCAGAAAGCTCC 948

Db 481 TTGATGCCGATTAACAAGATCGCTTTAAGGTTACGACGAAGCGCTCAGAAAGCTCC 540

QY 949 AGGTGTGGCCGTGGGACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAG 1008

Db 541 AGGTGTGGCCGTGGGACGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAG 600

QY 1009 GACACTGTCTCAAGCATCTTAGGCGCTTATGAAGAACTTAACAAGACAGGGGAC 1068

Db 601 GACACTGTCTCAAGCATCTTAGGCGCTTATGAAGAACTTAACAAGACAGGGGAC 660

QY 1069 CTGCCACCCCGACCAACCCGAGAGGGTGTGGCATCCAGGCTGTGCCGCTTCA 1128

Db 661 CTGCCACCCCGACCAACCCGAGAGGGTGTGGCATCCAGGCTGTGCCGCTTCA 720

QY 1129 CGTGGCGGGAAGACCTCTGCAAGGGGGAAGGCAAGGCCCTCCAGGCCACAGCCCGG 1188

Db 721 CGTGGCGGGAAGACCTCTGCAAGGGGGAAGGCAAGGCCCTCCAGGCCACAGCCCGG 779

QY 1189 TTCCCGCGGCCATTGTGTCAGCTTCCCGCCACGGGCACTGTTCTTCCACACCTG- 1247

Db 780 TTCCCGCGGCCATTGTGTCAGCTTCCCGCCACGGGCACTGTTCTTCCACACCTG- 839

QY 1248 TCCTGTGGGGCTGTCC-GGGAAGACACCTA-CCCTGTGGGCACTCAGGGTGTGCCAGC 1305

Db 840 TCCTGTGGGGCTGTCCGGGAAGACACCTA-CCCTGTGGGCACTCAGGGTGTGCCAGC 899

QY 1306 CCGG 1309

Db 900 CCGG 903

RESULT 5

LOCUS BI753688 909 bp mRNA linear EST 25-SEP-2001

DEFINITION 603023590F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5193959 5', mRNA sequence.

ACCESSION B1753688

VERSION B1753688.1 GI:15745266

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 909)

TITLE NIH-MGC http://mgs.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L14M1485 row: d column: 24
High quality sequence stop: 884.
Location/Qualifiers
1..909

FEATURES
Source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5193959"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 41.5%; Score 814.6; DB 3; Length 909;
Best Local Similarity 97.1%; Pred. No. 6.4e-207;
Matches 883; Conservative 0; Mismatches 19; Indels 7; Gaps 5;

QY 164 CGGAGCTGCTCTCTGACTGCGCGGAGCGGAACAATTGCGGGGAGATCAGCCGAA 223
|||||
Db 1 CGGAGCTGCTCTCTGACTGCGCGGAGCGGAACAATTGCGGGGAGATCAGCCGAA 60

QY 224 -AGAGCAAGTGGTGATATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTC 282
|||||
Db 61 GAGAGCAAGTGGTGATATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTC 120

QY 283 ATAGATCGAGCGTCAAGGGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTC 342
|||||
Db 121 ATAGATCGAGCGTCAAGGGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCTC 180

QY 343 CTGAACATTGAGGAATGAAGTTGAAAAACCCCGAAGATACCGATCATGAAGGAGAG 402
|||||
Db 181 CTGAACACTGAGGAATGAAGTTGAAAAACCCCGAAGATACCGATCATGAAGGAGAG 240

QY 403 GGCAAGAGTCACTTGAGCACATCCAGCGATCGACCCGGGACGTAAGCGGACATTAAAG 462
|||||
Db 241 GGCAAGAGTCACTTGAGCACATCCAGCGATCGACCCGGGACGTAAGCGGACATTAAAG 300

QY 463 AAGCATATATTCTTCAAGGATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTC 522
|||||
Db 301 AAGCATATATTCTTCAAGGATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTC 360

QY 523 CTGGCATATGAGGAGTATAACCCGGAGTGGGCTACTGACAGGACCTGAGCCACATCGCC 582

Db 361 CTGGCATATGAGGAGTATAACCCGGAGTGGGCTACTGACAGGACCTGAGCCACATCGCC 420
|||||

QY 583 GCCTGTTCCTCCTCTATCTTCTGAGGAGATGATCTGGGCACTGTGTGACGTCTG 642
|||||

Db 421 GCCTGTTCCTCCTCTATCTTCTGAGGAGATGATCTGGGCACTGTGTGACGTCTG 480
|||||

QY 643 GCCAGTGAGGCACTCCCTGACAGGATTTTCACAGCCCAAATGGCGGACCGTCCAGGGG 702
|||||

Db 481 GCCAGTGAGGCACTCCCTGACAGGATTTTCACAGCCCAAATGGCGGACCGTCCAGGGG 540
|||||

QY 703 CTCCAAGACCAACAGGAGCATGTGTAGCCACGTCAACAACCAAGACCATGGGATCAG 762
|||||

Db 541 CTCCAAGACCAACAGGAGCATGTGTAGCCACGTCAACAACCAAGACCATGGGATCAG 600
|||||

QY 763 GACAAGAAAGATCTATGTGGGCACTGTCCCGTTAGGCTGCTCATCCGATATTGATT 822
|||||

Db 601 GACAAGAAAGATCTATGTGGGCACTGTCCCGTTAGGCTGCTCATCCGATATTGATT 660
|||||

QY 823 GACGGGATCTCTCGGGCTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAGCGGAA 882
|||||

Db 661 GACGGGATCTCTCGGGCTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAGCGGAA 720
|||||

QY 883 CAGCGTGTATGCCG-ATAACAAGATCGCCTTAAGTTACAGCAGAAAGCGCTCACGAA 941
|||||

Db 721 CAGCGTGTATGCCGATTAACAAGATCGCCTTAAGTTACAGCAGAAAGCGCTCACGAA 780
|||||

QY 942 GACGTCCA-GTGTGGCCCGTGGG--CACGTTTTCGAACCGGTTGTTGATACCTGGGC 998
|||||

Db 781 GACGTTACGSGTGTGGCCCGTGGGCGCACGTTTTCGAACCGGTTGTTGATACCTGGGG 840
|||||

QY 999 CAGGG--ATGAGGACACTGTGTCAAGCATCTTAGGGCTCTATGAGAACTTAACAAG 1056
|||||

Db 841 CCAGGCATTGAGGACACTGTGTCAAGCATCTTAGGGCTCTATGAGAACTTAACAAG 900
|||||

QY 1057 AAGCAGGGG 1065
|||||

Db 901 AAGCAGGGG 909
|||||

RESULT 6
BU526609 954 bp mRNA linear EST 13-SEP-2002

LOCUS

DEFINITION AGENCOURT_10181753 NIH_MGC_101 Homo sapiens cDNA clone

IMAGE:6536307 5', mRNA sequence.

ACCESSION BU526609

VERSION BU526609.1 GI:22837050

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 954)

TITLE NIH-MGC http://mgs.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L14M2697 row: h column: 03
High quality sequence stop: 627.
Location/Qualifiers
1..954
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:6536307"
/cissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_101"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC library."

ORIGIN

Query Match	40.6%;	Score 797.6;	DB 5;	Length 954;
Best Local Similarity	98.1%;	Pred. No. 2.4e-202;		
Matches	828;	Conservative	0;	Mismatches 14; Indels 2; Gaps 2;
QY	517	ATCCTCTGGCAATATGAGAGTATAACCCGAGTGGGCTACTGACGGACCTGAGCCAC	576	
DB	1	ATCCTCTGGCAATATGAGAGTATAACCCGAGTGGGCTACTGACGGACCTGAGCCAC	60	
QY	577	ATCGCCGCTGTTCCTCTCTATCTTCTGTAGAGAGATGCATTCTGGGCACTGTGTGAC	636	
DB	61	ATCGCCGCTGTTCCTCTCTATCTTCTGTAGAGAGATGCATTCTGGGCACTGTGTGAC	120	
QY	637	CTGCTGGCCAGTGAAGGCACTCCCTGACGGGATTTCACAGCCCAATGGCGGACCGTC	696	
DB	121	CTGCTGGCCAGTGAAGGCACTCCCTGACGGGATTTCACAGCCCAATGGCGGACCGTC	180	
QY	697	CAGGGGCTCCAAGACCAACAGAGCATGTGTAGCCACGTCAACAACCAAGACCATGGG	756	
DB	181	CAGGGGCTCCAAGACCAACAGAGCATGTGTAGCCACGTCAACAACCAAGACCATGGG	240	
QY	757	CATCAGACAAGAAAGATCTATGTGGGCAGTTCGCCGTAGGCTGCCTCATCCGAGTA	816	
DB	241	CATCAGACAAGAAAGATCTATGTGGGCAGTTCGCCGTAGGCTGCCTCATCCGAGTA	300	
QY	817	TTGATTGACGGGATCTCTCTCGGGCTCACCCCTGCGCCTGTGGAAGTGTATCTGTAGAA	876	
DB	301	TTGATTGACGGGATCTCTCTCGGGCTCACCCCTGCGCCTGTGGAAGTGTATCTGTAGAA	360	
QY	877	GGCGAACAGGCGGTGATGCCGATAACAAGATCGCCTTAAAGTTTCAAGCAAGCGCCTC	936	
DB	361	GGCGAACAGGCGGTGATGCCGATAACAAGATCGCCTTAAAGTTTCAAGCAAGCGCCTC	420	
QY	937	ACGAAGACGTCAAGTGTGGCCCGTGGGCACTTTTTCACACCGGTTCTGTGATACCTGG	996	
DB	421	ACGAAGACGTCAAGTGTGGCCCGTGGGCACTTTTTCACACCGGTTCTGTGATACCTGG	480	
QY	997	GCCAGGATGAGACACTGTGCTCAAGCATTTAGGACCTCTATGAAGAACTAACAGA	1056	
DB	481	GCCAGGATGAGACACTGTGCTCAAGCATTTAGGACCTCTATGAAGAACTAACAGA	540	
QY	1057	AAGCAGGGGACCTTGCCACCCCGACCAAACCCGAGCAAGGTCGTCCGCATCCAGGCT	1116	
DB	541	AAGCAGGGGACCTTGCCACCCCGACCAAACCCGAGCAAGGTCGTCCGCATCCAGGCT	600	
QY	1117	GTGCCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGAGACAGGCCCCCTCCAGGC	1176	
DB	601	GTGCCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGAGACAGGCCCCCTCCAGGC	660	
QY	1177	CCACCAGCCCGGTTCCCGCGCCCATTTGCTCAGCTTCCCGCCACGGGCACTCGTTCT	1236	
DB	661	CCACCAGCCCGGTTCCCGCGCCCATTTGCTCAGCTTCCCGCCACGGGCACTCGTTCT	720	
QY	1237	TCCACACCCCTGTCTGTGGGCTGTCCGGGAAGACACCTACCTGTGGGCACTCAGGCT	1296	
DB	721	TTCACACCCCTGTCTGTGGGCTGTCCGGGAAGACACCTACCTGTGGGCACTCAGGCT	780	
QY	1297	GTGCCAGCCCGG--CCCTGGCTCAGGAGG--ACCTCAGGGTTCTTGAAGATTCCCTGAGT	1354	
DB	781	GTGCCAGCCCGGCCCCCTGGCTCAAGAGAGAACTCAGGGTTCTTGGGAAGATCCTGGCAG	840	

QY 1355 GGAA 1358
DB 841 GGGA 844

RESULT 7
BI522600
LOCUS
DEFINITION
603175612F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240090 5',
mRNA sequence.
BI522600
ACCESSION
BI522600.1 GI:15347392
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 919)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM11605 row: 9 column: 03
High quality sequence stop: 834.
Location/Qualifiers
1. 919

FEATURES
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5240090"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC library."

ORIGIN

Query Match	40.6%;	Score 797.4;	DB 3;	Length 919;
Best Local Similarity	96.1%;	Pred. No. 2.7e-202;		
Matches	872;	Conservative	0;	Mismatches 26; Indels 9; Gaps 5;
QY	895	CCGATAACAAGAAATCGCCTTTAAGTTTCAGCAGAAAGCGCTTCAGAAAGCTCCAGGTGT	954	
DB	1	CCGATAACAAGAAATCGCCTTTAAGTTTCAGCAGAAAGCGCTTCAGAAAGCTCCAGGTGT	60	
QY	955	GGCCCGTGGGCACTTTTTCACACCGGTTGTTGATACCTGGGCCAGGATGAGACACT	1014	
DB	61	GGCCCGTGGGCACTTTTTCACACCGGTTGTTGATACCTGGGCCAGGATGAGACACT	120	
QY	1015	GTGCTCAAGCATTTAGGCTCTTATGAAGAACTAAACAAGAAAGCAGGGGAGACTGCCA	1074	
DB	121	GTGCTCAAGCATTTAGGCTCTTATGAAGAACTAAACAAGAAAGCAGGGGAGACTGCCA	180	
QY	1075	CCCCCAGCCAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAGGTGCC	1134	
DB	181	CCCCCAGCCAACCCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAGGTGCC	240	

QY 1135 GGGAGACCTCTGCAAGGGGGACAGGAGGCCCCCTCCAGGCCCCACGACCCGGTTCCCG 1194
|||||
Db 241 GGGAGACCTCTGCAAGGGGGACAGGAGGCCCCCTCCAGGCCCCACGACCCGGTTCCCG 300
QY 1195 CGGCCCATTTGGTCACTTCCCGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGGT 1254
|||||
Db 301 CGGCCCATTTGGTCACTTCCCGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGGT 360
QY 1255 GGGGCTGTCCGGGAGACACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCCCTG 1314
|||||
Db 361 GGGGCTGTCCGGGAGACACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCCCTG 420
QY 1315 GCTCAGGAGGAGCTCAGGGTTCCTGAGATTCTTGAGTGAAGTCCATGCCCCGCTTC 1374
|||||
Db 421 GCTCAGGAGGAGCTCAGGGTTCCTGAGATTCTTGAGTGAAGTCCATGCCCCGCTTC 480
QY 1375 CCAACGAGCTGAGCTAGAGGGCCCTTGTTCCGCCATTATGATTTAGACAGAGCTGC 1434
|||||
Db 481 CCAACGAGCTGAGCTAGAGGGCCCTTGTTCCGCCATTATGATTTAGACAGAGCTGC 540
QY 1435 TGGGTCCGTGCAATCCCAAGAGACCACTGGCCCCCTGCTGGCAGGCTGAACACCT 1494
|||||
Db 541 TGGGTCCGTGCAATCCCAAGAGACCACTGGCCCCCTGCTGGCAGGCTGAACACCT 600
QY 1495 GCGGAGCGGGTGAATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCC 1553
|||||
Db 601 GCGGAGCGGGTGAATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGCAACCC 660
QY 1554 CTTCAGAGCTAGGAGCAAGACAGCCGTGTGCTCCCACTCAGGGCCTTGCTGCGGCT 1613
|||||
Db 661 CTTCAGAGCTAGGAGCAAGACAGCCGTGTGCTCCCACTCAGGGCCTTGCTGCGG-CT 719
QY 1614 CCACTTGGAAAGTTCTCAGTTCCTCCAGGCTTCTAGAA----GCATCTGGGCCAGGGC 1668
|||||
Db 720 CCACTTGGAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAAGATTTGGGCCACGGCT 779
QY 1669 TCATGCTGG-ATAATTTCCCTAGGCTTAACAACCCAGCAAGCTTC-GCGTCCCTGTTT 1726
|||||
Db 780 CCATGGCTGGCAATAATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGGCATCCTCGTT 839
QY 1727 TATTTTGGTTAACTTATGAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATTCA 1786
|||||
Db 840 TATGTATGTTAACTTATGAACTGTCTTAAGAAAGAGTGCCGTTGAGAGAGATTCCG 899
QY 1787 AGATGGA 1793
|||||
Db 900 AGATGGA 906

RESULT 8
CD653466 868 bp mRNA linear EST 18-JUN-2003
LOCUS
DEFINITION AGENCOURT_14539069 NIA Human H1 Embryonic Stem Cell cDNA Library
(Long) Homo sapiens cDNA clone IMAGE:30423486 5', mRNA sequence.
ACCESSION CD653466
VERSION CD653466.1 GI:31891804
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA

CDNA Library Preparation: Yulan Piao and Minoru Ko
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC c
Clone distribution information
can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: NDAM504 row: j column: 07
High quality sequence stop: 695.
Location/Qualifiers

FEATURES
source
1. 868
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30423486"
/issue_type="Embryonic Stem cells"
/cell_line="WA01"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIA Human H1 Embryonic Stem Cell cDNA library (long)"
/note="Vector: PCMV-Sport6; Site 1: NotI; Site 2: SalI;
This is a long-transcript enriched cDNA library (Genome
Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
cell line . Undifferentiated human ES cell line WA01/H1
was obtained from WCell Research Institute, Inc.,
Madison, WI, cultured according to their instructions, on
MEF feeders. They formed colonies with defined edges
and were positive for alkaline phosphatase, SSEA-4, OCT3,
OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days
after plating), the ES cells from 4 X 6cm dishes were
treated with 1 mg/ml collagenase, type IV
(Invitrogen/GIBCO) for 5-10 min and gently scraped off
with 5 ml pipette. RNA was purified with TRIzol Reagent
from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
(2001). [PMID:11544199]) Double-stranded cDNAs were
synthesized with an Oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGATCGCGAGCGCGCCCTTTTCTTTTCTT-3'] from
3.4g of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker L1-Sal4, purified by phenol/chloroform
extraction, and separated from free linkers by
Centricon-100 column. Then, the cDNAs were amplified by
long-range high fidelity PCR using Ex Taq polymerase
(Takara) with a primer Sal4-S for 25 cycles. The products
were purified by phenol/chloroform extraction and
Centricon-100 column. The cDNAs were digested with SalI
and NotI enzymes and cloned into SalI/NotI site of
PCMV-SPORT6 plasmid vector. The average insert size is
about 3.6kb."

ORIGIN

Query Match 39.9%; Score 783.6; DB 6; Length 868;
Best Local Similarity 97.6%; Pred. No. 1.3e-198;
Matches 803; Conservative 0; Mismatches 19; Indels 1; Gaps 1;
QY 448 AGCGGACATTAAGGAAGCATATATTTCTCAGGGATCGATACGGAACCAAGCAGCGGAA 507
|||||
Db 1 AGAGGTAATCCAAGGAAGCATATATTTCTCAGGGATCGATACGGAACCAAGCAGCGGAA 60
QY 508 CTACTCCACATCTCTCTGGCATATGAGAGTATAACCCGAGGTGGGCTACTGCAGGAC 567
|||||
Db 61 CTACTCCACATCTCTCTGGCATATGAGAGTATAACCCGAGGTGGGCTACTGCAGGAC 120
QY 568 CTGAGCCACATCGCCGCTGTGTTCTCTCTATCTTCTGAGAGAGATGATTCGGCA 627
|||||
Db 121 CTGAGCCACATCGCCGCTGTGTTCTCTCTATCTTCTGAGAGAGATGATTCGGCA 180
QY 628 CTGTGACAGTCTGGCCAGTGAAGGCACTCCCTGCAGGATTTACAGCCCAATGGC 687
|||||
Db 181 CTGTGACAGTCTGGCCAGTGAAGGCACTCCCTGCAGGATTTACAGCCCAATGGC 240
QY 688 GGGACCGTCCAGGGGCTCCAAGACCAAGAGCATGTGTGTAACCAAGTCAACCAAG 747

Db 241 GGGACCGTCCAGGGCTCCAAAGACCAAGAGCATGTGTAGCAAGTCAACAACCAAG 300
QY 748 ACCATGGGCGATCAGGACAAGAAAGATCTATGTGGCGAGTGTCCCGTTAGGCTGCTC 807
Db 301 ACCATGGGCGATCAGGACAAGAAAGATCTATGTGGCGAGTGTCCCGTTAGGCTGCTC 360
QY 808 ATCCGATATGTATGACGGGATCTCTCTCGGCTCACCCCTGCGCTGTGGAGCGTGTAT 867
Db 361 ATCCGATATGTATGACGGGATCTCTCTCGGCTCACCCCTGCGCTGTGGAGCGTGTAT 420
QY 868 CTGTTAGGAGGCGAAGCGGCTGTGATGCCGATTAACAAGATCGCTTTAAGGTTCAAGC 927
Db 421 CTGTTAGGAGGCGAAGCGGCTGTGATGCCGATTAACAAGATCGCTTTAAGGTTCAAGC 480
QY 928 AAGCGCTCAGGAGACGTCAGGTGTGGCCCGTGGGACGTTTGTCAACCGGTTGCTT 987
Db 481 AAGCGCTCAGGAGACGTCAGGTGTGGCCCGTGGGACGTTTGTCAACCGGTTGCTT 540
QY 988 GATACCTGGGCGAGGATGAGGACACTGTGTCAAGCATCTTAGGCTCTATGAAGAA 1047
Db 541 GATACCTGGGCGAGGATGAGGACACTGTGTCAAGCATCTTAGGCTCTATGAAGAA 600
QY 1048 CTACAAAGAAAGCAGGGGGACCTGCCACCCCCAGCCAAACCCGAGAGGGTCTCGGCA 1107
Db 601 CTACAAAGAAAGCAGGGGGACCTGCCACCCCCAGCCAAACCCGAGAGGGTCTCGGCA 660
QY 1108 TCCAGGCTGTGCGGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGAGCAGGAGCC 1167
Db 661 TCCAGGCTGTGCGGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGAGCAGGAGCC 720
QY 1168 CCTCAGGCGCCACGACCCCGGTTCCCGGCCCCATTGTCAGCTT-CCCCGCGACGGG 1226
Db 721 CCTCAGGCGCCACGACCCCGGTTCCCGGCCCCATTGTCAGCTTCCCCGCGACGGG 780
QY 1227 ACCTGTTCTTCCACACCTGTCTGTGGGGCTGTCCGGGA 1269
Db 781 ACCTGTTCTTCCACACCTGTCTGTGGGGCTGTCCGGGA 823

RESULT 9
BUS42901 881 bp mRNA linear EST 13-SEP-2002
LOCUS AGENCOURT 10334768 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574864
DEFINITION 5', mRNA Sequence.
ACCESSION BUS42901
VERSION BUS42901.1 GI:22853384
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 881)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LINC2770 row: n column: 16
High quality sequence stop: 642.
Location/Qualifiers
1 . 881
FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:6574864"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_40"
/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

Query Match 38.3%; Score 751.6; DB 5; Length 881;
Best Local Similarity 95.1%; Pred. No. 5.2e-190;
Matches 808; Conservative 0; Mismatches 39; Indels 3; Gaps 3;

QY 242 TGCTGGAGACTGGGAGAAATACAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 301
Db 2 TGCTGGAGACTGGGAG-AAATACAAAGCAGAGAAAGCTCATAGATCGAGCGTACAAG 60
QY 302 GAATGCCATGAACATCCGGGCGCCGATGTGTGTCAGTCCCTCTGAACATTGAGGAATGA 361
Db 61 GAATGCCATGAACATCCGGGCGCCGATGTGTGTCAGTCCCTCTGAACACTGAGGAATGA 120
QY 362 AGTTGAAAACCCCGGAAGATACCAATCATGAAGAGAAAGGCGAAGGTCATCTGAGC 421
Db 121 AGTTGAAAACCCCGGAAGATACCAATCATGAAGAGAAAGGCGAAGGTCATCTGAGC 180
QY 422 ACATCCAGCGCATCGACCGGGACGTAAAGCGGGACATTAAAGAGCATATATTCTTCAGGG 481
Db 181 ACATCCAGCGCATCGACCGGGACGTAAAGCGGGACATTAAAGAGCATATATTCTTCAGGG 240
QY 482 ATGATACGGAACCAAGCAGCGGGAATACTCCACATCTCTGCGCATATGAGAGATA 541
Db 241 ATGATACGGAACCAAGCAGCGGGAATACTCCACATCTCTGCGCATATGAGAGATA 300
QY 542 ACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCGGCTTCTCTCTATC 601
Db 301 ACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATCGCGGCTTCTCTCTATC 360
QY 602 TTCCTGAGGAGATGCATTCTGGGCACTGTGTCAGCTGTGCGGCAATGAGGCACTCCC 661
Db 361 TTCCTGAGGAGATGCATTCTGGGCACTGTGTCAGCTGTGCGGCAATGAGGCACTCCC 420
QY 662 TGCAGGGAATTCACAGCCCAATGGCGGGAACGCTCCAGGGGCTCCAAGACCAAGAGAGC 721
Db 421 TGCAGGGAATTCACAGCCCAATGGCGGGAACGCTCCAGGGGCTCCAAGACCAAGAGAGC 480
QY 722 ATGTGTAGCCACGTCAACAACCAAGACCATGAGGCGATCAGCAAGAAAGATCTATG 781
Db 481 ATGTGTAGCCACGTCAACAACCAAGACCATGAGGCGATCAGCAAGAAAGATCTATG 540
QY 782 GGCAGTGTCCCGGTTAGGCTGCTCATCCGATATTTGACGGGATCTCTCGGCG 841
Db 541 GGCAGTGTCCCGGTTAGGCTGCTCATCCGATATTTGACGGGATCTCTCGGCG 600
QY 842 TCACCCCTGGGCTGTGGACGTTATCTGTTAGAAAGGCGAAGCGGTTGATGCCGATA 901
Db 601 TCACCCCTGGGCTGTGGACGTTATCTGTTAGAAAGGCGAAGCGGTTGATGCCGATA 660
QY 902 CAAGAAATCGCTTTAAGGTTACAGAGAGCGCTCACAGAGCGTCCAGGTGTGGCCGT 961
Db 661 CAAGAAATCGCTTTAAGGTTACAGAGAGCGCTCACAGAGCGTCCAGGTGTGGCCGT 720
QY 962 GGGCAGTTTTTTCGAACCGGTTGCTGATACCT-GGGCAGGAGTGAAGACACT-GTGCT 1019
Db 721 GGGCAGTTTTTTCGAACCGGTTGCTGATACCTGGGCGGAGATGAAGACACTGTGCT 780
QY 1020 CAAGCATTTAAGGCTCTATGAAGAACTAACAAAGAGAGGAGGACCTGCCACCCC 1079
Db 781 CAAGCATTTAAGGCTCTCTATGAAGAACTAACAAAGAGGAGGAGGAGGAGGCC 840

QY 1080 AGCCAACCC 1089
| | | | |
Db 841 TGGCAAAACC 850

RESULT 10
BI597531 818 bp mRNA linear EST 07-SEP-2001
LOCUS 603247059F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5294607 5',
DEFINITION mRNA sequence.

ACCESSION BI597531
VERSION BI597531.1 GI:15490470
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 818)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palcovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1AM1745 row: f column: 16
High quality sequence stop: 811.

FEATURES
Source Location/Qualifiers
1..818
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5294607"
/issue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
Query Match 37.7%; Score 741.2; DB 3; Length 818;
Best Local Similarity 95.6%; Pred. No. 3.1e-187;
Matches 783; Conservative 0; Mismatches 34; Indels 2; Gaps 2;

QY 991 ACCTGGCCAGGATGAGACACTGTGCTCAGCATCTTAGGGCTCTATGAAGAACTA 1050
| | | | |
Db 1 AGCGGGTCAAGATGAGACACTGTGCTCAGCATCTTAGGGCTCTATGAAGAACTA 60

QY 1051 ACAAGAAAGCAGGGGACCTGCCACCCCAAGCCCAAGAGAGAGGTGTCGGCATCC 1110
| | | | |
Db 61 ACAAGAAAGAGGGGACCTGCCACCCCAAGCCCAAGAGAGAGGTGTCGGCATCC 120

QY 1111 AGGCTGTGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGGACAGGAGGCCCT 1170
| | | | |
Db 121 AGGCTGTGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGGACAGGAGGCCCT 180

QY 1171 CCAGGCCCAAGCCCGGTCCCGGCCCATTTGGTCAGCTTCCCGCCACGGGCACCT 1230

Db 181 CCAGGCTACCAAGCCCGGTTCCCGGGCCCATTTGGTCAGCTTCCGTGCACGGGCACCT 240
| | | | |
QY 1231 CGTCTTCCACACCCCTGTCTGTGTTGGGCTGTCCGGAGACACCTACCCTGTGGCACT 1290
| | | | |
Db 241 CGTCTTCCACACCCCTGTCTGTGTTGGGCTGTCCGGAGACACCTACCCTGTGGCACT 300
| | | | |

QY 1291 CAGGTTGCCCCAGCCCGGCTGTGCTCAGGAGGACCTCAGGGTTCTGAGATTCTCTG 1350
| | | | |
Db 301 CAGGTTGCCCCAGCCCGGCTGTGCTCAGGAGGACCTCAGGGTTCTGAGATTCTCTG 360
| | | | |

QY 1351 CAGTGAATCTCATGCCCCGCTCCCAAGGACCTGAGCTAGAGGGCCCTTGTTCGCC 1410
| | | | |
Db 361 CAGTGAATCTCATGCCCCGCTCCCAAGGACCTGAGCTAGAGGGCCCTTGTTCGCC 420
| | | | |

QY 1411 CATTATGATTTACAGACAGAGCTGTGGGTCCTGACATATCCAGAGAGACAGCTGGCC 1470
| | | | |
Db 421 CATTATGATTTACAGACAGAGCTGTGGGTCCTGACATATCCAGAGAGACAGCTGGCC 480
| | | | |

QY 1471 CCCTGCTGAGAGCTGAACACCTGCGGAGCGGGTGAATCGGCTTTCGTGACCCAGC 1530
| | | | |
Db 481 CCCTGCTGAGAGCTGAACACCTGCGGAGCGGGTGAATCGGCTTTCGTGACCCAGC 540
| | | | |

QY 1531 ACTGATTCGACCAAGGACACCCCTTCAGAGCTAGGAGCAACAGCCGTGTGCTCCACC 1590
| | | | |
Db 541 ACTGATTCGACCAAGGACACCCCTTCAGAGCTAGGAGCAACAGAGTGTGCTCCACC 600
| | | | |

QY 1591 TCAGGGCTGTGCTCTGCGGCTCCACTTGAAAGTTCTCAGTTCCTCCAGGCTTAG 1650
| | | | |
Db 601 TCAGGGCTGTGCTCTGCGG-CTCCACTTGAAAGTTCTCAGTTCCTCCAGGCTTAG 659
| | | | |

QY 1651 AAGCATCTGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAAAGCAAG 1710
| | | | |
Db 660 AAGCATCTGGCCAGGGCTCATGGCTGATTAATTTCCCTAGGCTTAACAACCCAAAGCAAG 719
| | | | |

QY 1711 CTTGCGTCTCTGTTTA-TTTTGTGTTAACTTATGAATAATGTTAAGAAAGAGTGA 1769
| | | | |
Db 720 CTTGCGATCCTCGTNTTATTTTGTGTTAACTTATGAATAATGTTAAGAAAGAGTGC 779
| | | | |

QY 1770 GCTCGAGAGATTCAGAGATGAACACACAGACCCCA 1808
| | | | |
Db 780 AGTCAGAGAAATTCCGAGATGAACACACAGAACCA 818
| | | | |

RESULT 11
BQ920917 930 bp mRNA linear EST 20-AUG-2002
LOCUS BQ920917
DEFINITION AGENCOURT 8926259 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6462826
5', mRNA Sequence.

ACCESSION BQ920917
VERSION BQ920917.1 GI:22335834
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE 1 (bases 1 to 930)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1CM2653 row: f column: 11
High quality sequence stop: 658.

FEATURES
Location/Qualifiers

```
source
1. .930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6462826"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 101"
/note="Organ: lung; Vector: pOTB7; Site_1: EcoRI; Site_2:
XhoI; cDNA made by oligo-dt priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match      37.3%; Score 732.6; DB 5; Length 930;
Best Local Similarity 97.2%; Pred. No. 6.6e-185;
Matches 810; Conservative 0; Mismatches 14; Indels 9; Gaps 6;

QY 482 ATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGCGCATATGAGAGATATA 541
    |||
Db 1 ATCGATACGGAACC-AGCAGCGGAACTACTCCACATCCTCTGCG-ATATGAGAG-ATA 57

QY 542 ACCCGAGGTGGGCTACTGCAAGGAGCTGAGCCACATCGCGCCTTGTTCCTCCTATC 601
    |||
Db 58 ACCCGAGGTGGGCTACTGCAAGGAGCTGAGCCACATCGCGCCTTGTTCCTCCTATC 117

QY 602 TTCTTGAGAGGATGATCTTGCGGACTGTGTCAGCTGTGCGCCAGTGAGAGGCACTCCC 661
    |||
Db 118 TTCTTGAGAGATG-ATTCTGGGACTGTGTCAGCTGTGCGCCAGTGAGAGGCACTCCC 176

QY 662 TGCAAGGATTTCAACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAGAGC 721
    |||
Db 177 TGCAAGGATTTCAACAGCCCAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAGAGC 236

QY 722 ATGTGTAGCAGCGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGTG 781
    |||
Db 237 ATGTGTAGCAGCGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGTG 296

QY 782 GGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGGC 841
    |||
Db 297 GGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTGATTGACGGGATCTCTCGGGC 356

QY 842 TCACCCCTGGGCTGTGGGACGTATCTGTGTAAGGCGGAACAGCGTTGATGCCATTA 901
    |||
Db 357 TCACCCCTGGGCTGTGGGACGTATCTGTGTAAGGCGGAACAGCGTTGATGCCATTA 416

QY 902 CAAGATGCGCTTTAAGGTTCAAGAAAGCGCTCAAGAAAGCGTCCAGGTGTGCCCGT 961
    |||
Db 417 CAAGATGCGCTTTAAGGTTCAAGAAAGCGCTCAAGAAAGCGTCCAGGTGTGCCCGT 476

QY 962 GGGCAGTTTTTGCAACCGGTTGTTGATACCTGGGCAAGGATGAGGACATGTGCTCA 1021
    |||
Db 477 GGGCAGTTTTTGCAACCGGTTGTTGATACCTGGGCAAGGATGAGGACATGTGCTCA 536

QY 1022 AGCATTTAAGGCTCTATGAAGAACTAACAAGAAAGCAGGGGACCTGCCACCCCAAG 1081
    |||
Db 537 AGCATTTAAGGCTCTATGAAGAACTAACAAGAAAGCAGGGGACCTGCCACCCCAAG 596

QY 1082 CCAAAACCGAGCAAGGGTCTGCGGATCCAGGCTGTGCGGGCTTCACGTGGCGGAAGA 1141
    |||
Db 597 CCAAAACCGAGCAAGGGTCTGCGGATCCAGGCTGTGCGGGCTTCACGTGGCGGAAGA 656

QY 1142 CCCTCTGAAGGGGACAGGAGCCCTCCAGGCCCAAGCCCGTTCCCGCGGCCCA 1201
    |||
Db 657 CCCTCTGAAGGGGACAGGAGCCCTCCAGGCCCAAGCCCGTTCCCGCGGCCCA 716

QY 1202 TTGTGTCAGCTTCCCGGCAAGGGGACCTCGTTCTTCAACCCCTGTCC--TGTTGGGGC 1259
    |||
Db 717 TTGTGTCAGCTTCCCGGCAAGGGGACCTCGTTCTTCAACCCCTGTCCCTGTGGGGCC 776
```

```
QY 1260 TGTCCGGGAAGACACCTACC--TGTTGGCACTCAGGGTGTGCCACGCCCG 1309
    |||
Db 777 TGTCCGGGAAGACACCTACCCTGTGGGAACCTCAGGGTGTGCCACGCCCG 829

RESULT 12
BMS60320
LOCUS
DEFINITION
BMS60320 1013 bp mRNA linear EST 20-FEB-2002
AGENCOURT_6564065 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5744726
5', mRNA sequence.
BMS60320
BMS60320.1 GI:18804652
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1013)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12766 row: 1 column: 15
High quality sequence start: 36
High quality sequence stop: 745.
location/Qualifiers
1. 1013
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744726"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dt primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match      37.2%; Score 730.2; DB 3; Length 1013;
Best Local Similarity 92.6%; Pred. No. 3e-184;
Matches 801; Conservative 0; Mismatches 58; Indels 6; Gaps 3;

QY 1 ATGACGTGTAGAGTTCGGGCAAGTTGTGGGCAAGAGCGAGAGACATCATATG 60
    |||
Db 98 ATGACGTGTAGAGTTCGTGGGTAGTTGTGGGCAAGAGCGAGAGACATCATATG 157

QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTAAGCTTTTGA 120
    |||
Db 158 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGGCTAAGCTTTTGA 217

QY 121 AGCTACACAACAACGTGATCATTTGGGATTGTATGATGAGACGAGCTGCTCTG 180
    |||
Db 218 AGCTACACAACAACGTGATCATTTGGGATTGTATGATGAGACGAGCTGCTCTG 277

QY 181 ACTCGCGGAGCGAAGCAATTCGGCGGAGATCAGCGAAAGAGCAAGTGGGTGAT 240
    |||
```

Db 278 ACTGCGGGAGGTGAAGCAATTTCGGCCGGAGATCAGCCGAAAGCAAGTGGTGAAA 337

Qy 241 ATGCTGGAGACTGGAGAAATACAAAAGCAGAAAAGCTCATAGATCGAGCGTACAG 300

Db 338 ATGCTGGAGAAATGGGACACCTACAAAACAGCAGAAAAGCTCATAGATCGAGCGTACCAG 397

Qy 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTCTGAACATTGAGAAATG 360

Db 398 GGAATTCCTCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTCTGAACATTGAGAAATC 457

Qy 361 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAGAGGCAAGAGTCACTGTAG 420

Db 458 AAGTTGAAAAACCCCGGAAGATACCAATCATGAAGAGAGAGGCAAGAGTCACTGTAG 517

Qy 421 CACATCCAGCGCATCGACCGGGACGTAAAGCGGACATTAAAGAACATATATTCTTCAGG 480

Db 518 CACATCCAGCAGATGGAACCTGGACGTAAAGTGGGACATTAAAGAGGCATATATTCTTCAGG 577

Qy 481 GATGATACGGAACCAAGCAGCGGGAATCTCCACATCTCTGGCATATGAGAGATAT 540

Db 578 GATGATACGGAACCAAGCAGCGGGAATCTTACATCTCTGGCATATGAGAGATAT 637

Qy 541 AACCCGAGGTGGGCTACTGCAGGGAACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600

Db 638 AACCCGAGGTGGGCTACTGCAGGGAACCTGAGCCACATCGCCGCTTGTCTCTCTAT 697

Qy 601 CTTCTGAGGAGATGATCTTCTGGGCACTGTGCAGCTGCTGGCCAGTGAGAGCACTCC 660

Db 698 CTTCTGAGGAGATGATCTTCTGGGCACTGTGCAGCTGCTGGCCAGTGAGAGCACTCC 757

Qy 661 CTGCAGGATTTCAACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAACAGAG 720

Db 758 CTGCAGGATTTCAACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAACAGAG 817

Qy 721 CATGTGTTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAATCTATGT 780

Db 818 CATGTGTTAGCCACGTCAACTACCAAGACCATGTGGCATCAGGACAAGAAATCTATGT 877

Qy 781 GGGCAGTGTTC--CCGTTAGGCTGCTCATCC--GGATATTGATGACGGG--ATCTCT 834

Db 878 GGGCAGTGTTCCTTCCTTAGGCTGCTCATCCGGATATTGATGATGGATCTCTCC 937

Qy 835 CTCGGGCTCACCCCTGCGCCTGTGG 859

Db 938 TCGGCTCACCCCTGCGCCTGTGG 962

RESULT 13

BI601698 839 bp mRNA linear EST 07-SEP-2001

LOCUS 603249940F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5301752 5',

DEFINITION mRNA sequence.

ACCESSION BI601698

VERSION BI601698.1 GI:15494637

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 839)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.lnl.gov

Plate: LLM1763 row: p column: 09

High quality sequence stop: 732.

Location/Qualifiers

1. .839

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5301752"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/clone_lib="NIH_MGC_96"

/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 36.9%; Score 724.8; DB 3; Length 839;

Best Local Similarity 94.5%; Pred. No. 8e-183;

Matches 794; Conservative 0; Mismatches 42; Indels 4; Gaps 4;

Qy 991 ACCTGGGCCAGGATGAGACACTGTGCTCAAGCATCTTAGGGCTCTATAGAAACTA 1050

Db 1 AGCGGGGTGAGCATGAGACACTGTGCTCAAGCATCTTAGGGCTCTATAGAAACTA 60

Qy 1051 ACAAGAACAGGGGGACCTGCCACCCCCAGCCAAACCCAGCAAGGTCGTGGCATCC 1110

Db 61 ACAAGAACAGGGGGACCTGCCACCCCCAGCCAAACCCAGCAAGAGTCGTGGCATCC 120

Qy 1111 AGGCTGTGGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGACAGGCGCCCT 1170

Db 121 AGGCTGTGGCGGCTTCAAGTGGCGGGAAGACCTCTGCAAGGGGACAGGCGCCCT 180

Qy 1171 CCAGGCCCAAGCCCGGTTCCCGCGGCCATTGTGTCAGCTTCCCGCACGGGCACT 1230

Db 181 CCAGGCCCAAGCCCGGTTCCCGCGGCCATTGTGTCAGCTTCCCGCACGGGCACT 240

Qy 1231 CGTTCTTCCACACCCTGTCTGTGGGGCTGTCCGGGAAGACACTACCCTGTGGCACT 1290

Db 241 CGTTCTTCCACACCCTGTCTGTGGGGCTGTCCGGGAAGACACTACCCTGTGGCACT 300

Qy 1291 CAGGTTGCCCCAGCCCGGCTTCCAGGAGGAGCACTCAGGGTCTCTGAGATTCTCTG 1350

Db 301 CAGGTTGCCCCAGCCCGGCTTCCAGGAGGAGCACTCAGGGTCTCTGAGATTCTCTG 360

Qy 1351 CAGTGAACCTCATGCCCCGCTTCCAAACGGAACCTGAGAGGGCCCTTGTGTTCCGC 1410

Db 361 CAGTGAACCTCATGCCCCGCTTCCAAACGGAACCTGAGAGGGCCCTTGTGTTCCGC 420

Qy 1411 CATTATGATTTCAGACAGAGCTGTGGGTCCGTGCCATATCCAGAGAGCAAGCTGGCC 1470

Db 421 CATTATGATTTCAGACAGAGCTGTGGGTCCGTGCCATATCAAGAGAGCAAGCTGGCC 480

Qy 1471 CCCTGTGCGAGGCTGAACACCCCTGCGGAGCGGGTGAGATCGGCTTGTGTCACCCAGC 1530

Db 481 CCCTGTGCGAGGCTGAACACCCCTGCGGAGCGGGTGAGATCGGCTTGTGTCACCCAGC 540

Qy 1531 ACTGATTCGACCAAGGACCCCTTCAGAGCTAAGGAGCAACAGCCGTGTGCTCCACC 1590

Db 541 ACTGATTCGACCAAGGACCCCTTCAGAGCTAAGGAGCAACAGAGTGTGCTCCACC 600

Qy 1591 TCAGGGCTTGCCTCTGCGGCTTCACTTGAAAGTTCTCAGTTCCTCCAGGCTTCTAG 1650

Db 601 TCAGGGCTTGCCTCTGCGGCTTCACTTGAAAGTTCTCAGTTCCTCCAGGCTTCTAG 660

Qy 1651 AAGCATCTGGGCGCAGGGCTCATGCTGATTAATTTCCCTTAGGCTTAACAACCCAGCAAG 1710

Db 661 AAGCATCTGGGCCAGGGCTCATGCGTGATTA-TTCCCTAGGCTTTAAACAACCCAGCAAG 719

Qy 1711 CTTCGGCTCTCGGTTTATTGTTAACTTATGAAATGTATTAAGAA-AGAGTGCA 1769

Db 720 CTTCGCATCTCGGTTTATGTGGTTAACCTTATGAAAAGTCTTTAAGAACAGAGTGCA 779

Qy 1770 GCTCGA-GAGAGATTGAGA-GATGAACACACAGACCCAGATCACAAAAGCCAAACCATG 1827

Db 780 GCTCGAGAGAGATTCCGAGATGAACAGACAGAACCCAGATCACAAAAGCAAGCCTG 839

RESULT 14

BM553146 1141 bp mRNA linear EST 20-FEB-2002

LOCUS BM553146

DEFINITION AGENCOURT_6542478 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742850

ACCESSION BM553146

VERSION BM553146.1 GI:18791621

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1141)

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LLM12761 row: k column: 11

High quality sequence stop: 665.

Location/Qualifiers

1..1141

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5742850"

/issue_type="medulla"

/lab_host="DH10B"

/clone_lib="NIH_MGC_119"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH_MGC library."

ORIGIN

Query Match 36.6%; Score 718.2; DB 3; Length 1141;

Best Local Similarity 97.5%; Pred. No. 5,1e-181;

Matches 740; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

Qy 1 ATGAGCTGTAGAGTTCGGCGGAGTTGGTGGCACACAGAGCGAGACATCATTTATG 60

Db 102 ATGAGCTGTAGAGTTCGGCGGAGTTGGTGGCACACAGAGCGAGACATCATTTATG 161

Qy 61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGGCTTAAGCCTTTTGA 120

Db 162 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAGGGGCTTAAGCCTTTTGA 221

Qy 121 AGCTACAACAACAGCTGATCATTTGGGATTTGACATGAGACGAGCTGCTCTCTG 180

Db 222 AGCTACAACAACAACGTCATCATTTGGGATTTGTACATGAGACGAGCTGCCCTCTTG 281

Qy 181 ACTGCGCGGAGCGGAAGCAAAATTCGGCGGAGATCAGCCGAAGACAAAGTGGTGAT 240

Db 282 ACTGCGCGGAGCGGAAGCAAAATTCGGCGGAGATCAGCCGAAGACAAAGTGGTGAT 341

Qy 241 ATGCTGGAGACTGGGAGAAATACAAAGCAGACGAAAGCTCATAGATCGAGCTACAAG 300

Db 342 ATGCTGGAGACTGGGAGAAATACAAAGCAGACGAAAGCTCATAGATCGAGCTACAAG 401

Qy 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGCTCACTCCTCTGAACATTGAGGAATG 360

Db 402 GGAATGCCCATGAACATCCGGGGCCCGATGTGCTCACTCCTCTGAACATTGAGGAATG 461

Qy 361 AAGTGAACAAACCCCGAAGATACCAATCATGAAGAGAGGCAAGAGTCATCTGAG 420

Db 462 AAGATGAACAAACCCCGAAGATACCAATCATGAAGAGAGGCAAGAGTCATCTGAG 521

Qy 421 CACATCCAGCGCATGACCGGGACGTAAGCGGACATTAAAGACATATATTCTTCAGG 480

Db 522 CACATCCAGCGCATGACCGGGACGTAAGCGGACATTAAAGACATATATTCTTCAGG 581

Qy 481 GATCGATACGGAACCAAGCAGCGGGAATCTACATCATCTCTGTCATATGAGAGTAT 540

Db 582 GATCGATACGGAACCAAGCAGCGGGAATCTACATCATCTCTGTCATATGAGAGTAT 641

Qy 541 AACCAGAGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600

Db 642 AACCAGAGTGGGCTACTGACAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 701

Qy 601 CTTCCTGAGAGATGATCATTTCTGGGCACTGTGCACTGCTGCGCACTG-AGAGGCACTC 659

Db 702 CTTCCTGAGAGATGATCATTTCTGGGCACTGTGCACTGCTGCGCACTGAGTAAAGGCACTTC 761

Qy 660 CTTGACAGGATTTACACAGCCCAATGGCGGAGCCGTCACAGGGGCTCCAAAGACCAAGAGA 719

Db 762 CTTGACAGGATTTACACAGCCCAATGGCGGAGCCGTCACAGGGGCTCCAAAGACCAAGAGA 821

Qy 720 GCATGTGTAGCCACGTCACAACCCCAAGACCATGGGCA 758

Db 822 GCATGTGTAGCCACGTCACAACCCCAAGACCATGGGCA 860

RESULT 15

BI828272 831 bp mRNA linear EST 04-OCT-2001

LOCUS BI828272

DEFINITION 603078163F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5169999 5', mRNA sequence.

ACCESSION BI828272

VERSION BI828272.1 GI:15939822

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 831)

TITLE NIH-MGC http://mgc.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LLM11422 row: n column: 16

High quality sequence stop: 822.

Location/Qualifiers

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2006, 13:18:55 ; Search time 9933 Seconds
(without alignments)
11239.358 Million cell updates/sec

Title: US-10-071-838-1

Perfect score: 1964

Sequence: 1 atggagcgtgtagagtcgc.....aaatgttcctgtgaatc 1964

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_cm:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_sts:*

11: gb_sy:*

12: gb_un:*

13: gb_vi:*

14: gb_htg:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1964	100.0	1964	6	AX775945	AX775945 Sequence
2	1964	100.0	1964	8	AF540953	AF540953 Homo sapi
3	1962.4	99.9	1964	6	AX775947	AX775947 Sequence
4	1952.8	99.4	2090	8	BC071680	BC071680 Homo sapi
5	1948	99.2	2072	6	AX086847	AX086847 Sequence
6	1936.4	98.6	2166	8	BC075809	BC075809 Homo sapi
7	1902	96.8	2304	6	AX775943	AX775943 Sequence
8	1900.4	96.8	2304	6	AX775941	AX775941 Sequence
9	1836.4	93.5	2084	6	AR578516	AR578516 Sequence
10	1812.4	92.3	2072	6	AR578515	AR578515 Sequence
11	1809	92.1	2647	6	AX775951	AX775951 Sequence
12	1807.4	92.0	2647	6	AX775949	AX775949 Sequence
13	1807.4	92.0	2647	6	AX833552	AX833552 Sequence
14	1807.4	92.0	2647	6	AX095385	AX095385 Homo sapi
15	1799.4	91.6	2111	8	BC058890	BC058890 Homo sapi
16	1753.6	89.3	3389	8	AK122833	AK122833 Homo sapi
17	1612.4	82.1	1924	8	BC078140	BC078140 Homo sapi
18	1471.8	74.9	2575	8	BC033670	BC033670 Homo sapi

19	1220	62.1	7878	6	AX332512	AX332512 Sequence
20	1220	62.1	7878	8	HSTRB210	X63546 H. sapiens m
21	1218.4	62.0	7971	8	AY143550	AY143550 Homo sapi
22	1213.6	61.8	8491	8	HSM807865	BX647719 Homo sapi
23	999	50.9	8201	6	I76205	I76205 Sequence 9
24	999	50.9	8201	6	AX774956	AX774956 Sequence
25	999	50.9	8201	8	HSTRB213	X63547 H. sapiens m
26	882.8	44.9	123925	8	AC003976	AC003976 Homo sapi
27	879.6	44.8	192881	8	AC067923	AC067923 Homo sapi
28	878	44.7	37757	8	AC139091	AC139091 Homo sapi
29	878	44.7	41266	14	AC139176	AC139176 Homo sapi
30	878	44.7	42073	14	AC144840	AC144840 Homo sapi
31	878	44.7	89147	14	AC146857	AC146857 Homo sapi
32	878	44.7	158698	14	AC036181	AC036181 Homo sapi
33	878	44.7	163628	14	AC139367	AC139367 Homo sapi
34	878	44.7	185449	14	AC139361	AC139361 Homo sapi
35	878	44.7	193646	8	AC131056	AC131056 Homo sapi
36	876.4	44.6	134506	14	AC146482	AC146482 Homo sapi
37	876.4	44.6	140152	8	AC027821	AC027821 Homo sapi
38	876.4	44.6	140152	8	AC027821	AC027821 Homo sapi
39	876.4	44.6	146243	14	AC130293	AC130293 Homo sapi
40	876.4	44.6	180876	8	AC021317	AC021317 Homo sapi
41	876.4	44.6	181517	8	AC126327	AC126327 Homo sapi
42	876.4	44.6	206040	14	AC132814	AC132814 Homo sapi
43	874.8	44.5	36160	14	AC139089	AC139089 Homo sapi
44	874.8	44.5	172770	14	AC023133	AC023133 Homo sapi
45	869.2	44.3	91225	14	AC139556	AC139556 Homo sapi

ALIGNMENTS

RESULT 1	AX775945	1964 bp	mRNA	linear	PAT 14-JUL-2003
LOCUS	AX775945	Sequence 215	From Patent WO03048202.		
DEFINITION	AX775945				
ACCESSION	AX775945				
VERSION	AX775945.1	GI:32693663			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 Matsuda, A. and Muramatsu, S.				
AUTHORS	NE-KB activating gene				
TITLE	Patent: WO 03048202-A 215 12-JUN-2003;				
JOURNAL	Asahi Kasei Kabushiki Kaisha (JP)				
FEATURES	Location/Qualifiers				
source	1. 1964				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	1. 1650				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAE11672.1"				
	/db_xref="GI:32693664"				
	/translation="MDVBEVAGSWAQRERDIIMKYEGHRAGLPEDKPKPFRSYNNVDHLGIVHETELPPLTAREAKQIRREISRSKSKVMDMLGDMETKSSRKLIDRAYKGM PMNIRGPMWSTVLINIEBMKLNPNRYQIMKEKGRSSSEHQRIDRDVSGTLRKHIFPR DRYGKQRELHILHLLAYEENPEVQCRDLSHIALFLYLDEDAFWALVQLASER HSLGTFHSPNGTGVQGLQDOQEHVATSPKTMGHODKDLGQCSSPLGCLIRILIDG ISLGLRLMDVYLVEGEQALMPTTRIAFKVQQRLLTKTSRCGPWFARFNRFDTWAR DEDTVLKHILASMKKLTTRKQGLDPPAKPEQSSASRPVPASRGKTLCKDRQAPPG PPARFPRTWSASPPRAPRSSTPCGAVREDTYPVGTGVSPALAOQGPQGSWFL QWNSMRLPTDLDVEGPFWRHYDFRQSCWRAISQEDQLAPCWQAEHPAERVRSAPAA PSTSDSQGTFRARDEQPCAPTSGPCLGLHLESSQFPGF"				
CDS					
ORIGIN					
Query Match	100.0%;	Score 1964;	DB 6;	Length 1964;	
Best Local Similarity	100.0%;	Pred. No. 0;			

Matches 1964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	ATGACGTGGTAGAGTTCGGGCACTGTGGGCACAAGACGAGAGACATCATATATG	60	
Db	1	ATGACGTGGTAGAGTTCGGGCACTGTGGGCACAAGACGAGAGACATCATATATG	60	
QY	61	AAATACGAAAAGGACACCGAGCTGGCTGCAGAGGACAAGGGCCTTAAGCCTTTTGA	120	
Db	61	AAATACGAAAAGGACACCGAGCTGGCTGCAGAGGACAAGGGCCTTAAGCCTTTTGA	120	
QY	121	AGCTACACAACAACGTCGATCATTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG	180	
Db	121	AGCTACACAACAACGTCGATCATTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG	180	
QY	181	ACTGCGGGGAGGCGAAGCAATTGCGCGGAGATCAGCCGAAAAGCAAGTGGTGAT	240	
Db	181	ACTGCGGGGAGGCGAAGCAATTGCGCGGAGATCAGCCGAAAAGCAAGTGGTGAT	240	
QY	241	ATGCTGGAGACTGGAGAAATACAAAAGCAGAGAAAGCTATGATCGAGCGTACAAG	300	
Db	241	ATGCTGGAGACTGGAGAAATACAAAAGCAGAGAAAGCTATGATCGAGCGTACAAG	300	
QY	301	GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG	360	
Db	301	GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG	360	
QY	361	AAATTGAAAAACCCCGGAAGATACGATCATGAAGAGAAAGGCAAGAGTCACTGAG	420	
Db	361	AAATTGAAAAACCCCGGAAGATACGATCATGAAGAGAAAGGCAAGAGTCACTGAG	420	
QY	421	CACATCCAGCCGATCGACCGGGACGTAAGCGGGACATTAAAGAACATATATTCTTCAGG	480	
Db	421	CACATCCAGCCGATCGACCGGGACGTAAGCGGGACATTAAAGAACATATATTCTTCAGG	480	
QY	481	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAT	540	
Db	481	GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCCTCTGGCATATGAGAGTAT	540	
QY	541	AACCCGGAGTGGGCTACTGACAGGACCTGAGCCACATCGCGCCTTGTCTCTCTAT	600	
Db	541	AACCCGGAGTGGGCTACTGACAGGACCTGAGCCACATCGCGCCTTGTCTCTCTAT	600	
QY	601	CTTCCTGAGAGATGATTCGTGGCACTGTGACAGCTGCTGGCCAAGTGAAGGCACTCC	660	
Db	601	CTTCCTGAGAGATGATTCGTGGCACTGTGACAGCTGCTGGCCAAGTGAAGGCACTCC	660	
QY	661	CTGCAAGGATTTCAACAGCCCAATGCGGGACCGCTCCAGGGCTCCAAGCCAAGAG	720	
Db	661	CTGCAAGGATTTCAACAGCCCAATGCGGGACCGCTCCAGGGCTCCAAGCCAAGAG	720	
QY	721	CATGTGTAGCCACGTCACAACCCAGAACCATTGGGCAATCAGGACAAGAAAGATCTATGT	780	
Db	721	CATGTGTAGCCACGTCACAACCCAGAACCATTGGGCAATCAGGACAAGAAAGATCTATGT	780	
QY	781	GGGCAGTGTTCCCGTTAGGCTGCTCATCCGATATTGATGACGGATCTCTCGGG	840	
Db	781	GGGCAGTGTTCCCGTTAGGCTGCTCATCCGATATTGATGACGGATCTCTCGGG	840	
QY	841	CTCACCCCTGCGCTGTGGACGTGTACTGTGTAAGGCGGAACAGCGTTGATGCCGATA	900	
Db	841	CTCACCCCTGCGCTGTGGACGTGTACTGTGTAAGGCGGAACAGCGTTGATGCCGATA	900	
QY	901	ACAAGAAATGCGCTTTAAGTTCAAGAGAAGCGCCTCAGAAAGCTCCAGGTGCCCCG	960	
Db	901	ACAAGAAATGCGCTTTAAGTTCAAGAGAAGCGCCTCAGAAAGCTCCAGGTGCCCCG	960	
QY	961	TGGGACGTTTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGAAGGACATGTGCTC	1020	
Db	961	TGGGACGTTTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGAAGGACATGTGCTC	1020	
QY	1021	AAGCATCTTAAGGCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA	1080	
Db	1021	AAGCATCTTAAGGCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA	1080	

QY	1081	GCCAAACCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG	1140	
Db	1081	GCCAAACCGAGCAAGGTCGTCCGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG	1140	
QY	1141	ACCCTTGCAAGGGGACAGGCAAGCCCCCTCCAGGCCCAACCAAGCCCCGTTCCCGGGCC	1200	
Db	1141	ACCCTTGCAAGGGGACAGGCAAGCCCCCTCCAGGCCCAACCAAGCCCCGTTCCCGGGCC	1200	
QY	1201	ATTGTGACGCTTCCCCGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGTGGGCT	1260	
Db	1201	ATTGTGACGCTTCCCCGCCACGGGCACTCGTTCTTCCACACCCCTGTCTGTGGGCT	1260	
QY	1261	GTCGGGAAGACACCTACCTGTGGCACTCAGGCTGCCAGCCCCGCTGGCTCAG	1320	
Db	1261	GTCGGGAAGACACCTACCTGTGGCACTCAGGCTGCCAGCCCCGCTGGCTCAG	1320	
QY	1321	GGAAGACCTCAGGTTCTTGAGATTCTGACGTGAACTCCATGCCCGCTCCCAAG	1380	
Db	1321	GGAAGACCTCAGGTTCTTGAGATTCTGACGTGAACTCCATGCCCGCTCCCAAG	1380	
QY	1381	GACCTGACGTAAGAGGCGCTTGCTCCGCATTATGATTTACAGACAGCTGCTGGTC	1440	
Db	1381	GACCTGACGTAAGAGGCGCTTGCTCCGCATTATGATTTACAGACAGCTGCTGGTC	1440	
QY	1441	CGTGCAATATCCAGAGAGACCAAGCTGGCCCCCTGTGCGCAGGCTGAACACCTCGGAG	1500	
Db	1441	CGTGCAATATCCAGAGAGACCAAGCTGGCCCCCTGTGCGCAGGCTGAACACCTCGGAG	1500	
QY	1501	CGGGTGAGATCGGCTTTCGTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA	1560	
Db	1501	CGGGTGAGATCGGCTTTCGTGCAACCCAGCACTGATTCGACCAAGGCAACCCCTTCAGA	1560	
QY	1561	GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGCGCTTGCTTGCGGCTCACTTG	1620	
Db	1561	GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGCGCTTGCTTGCGGCTCACTTG	1620	
QY	1621	GAAAGTTCTAGTCCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGCTCATGCTGGAT	1680	
Db	1621	GAAAGTTCTAGTCCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGCTCATGCTGGAT	1680	
QY	1681	AATTTCCCTAAGCTTAAACAACCCAAAGCAAGCTTCGGCTCTCTTTATTTTGGTTAA	1740	
Db	1681	AATTTCCCTAAGCTTAAACAACCCAAAGCAAGCTTCGGCTCTCTTTATTTTGGTTAA	1740	
QY	1741	CTTATGAATAATGTATTAGAAAGAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACACC	1800	
Db	1741	CTTATGAATAATGTATTAGAAAGAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACACC	1800	
QY	1801	AGACCCGATCACAAAAGCCAAACCATGCCAGGCCCTCCAGCAACCCCAAGCCCAAGAC	1860	
Db	1801	AGACCCGATCACAAAAGCCAAACCATGCCAGGCCCTCCAGCAACCCCAAGCCCAAGAC	1860	
QY	1861	CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG	1920	
Db	1861	CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG	1920	
QY	1921	ATAACTTACTTCAAGTTTGAATAATGTTTCTGTGAAATG	1964	
Db	1921	ATAACTTACTTCAAGTTTGAATAATGTTTCTGTGAAATG	1964	

RESULT 2	
AF540953	1964 bp mRNA linear PRI 02-OCT-2002
LOCUS	AF540953
DEFINITION	Homo sapiens Rab GTPase-activating protein PRC17 (PRC17) mRNA,
ACCESSION	AF540953
VERSION	AF540953.1 GI:23452664
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 1964)
TITLE	Pel, L., Peng, Y., Ling, X.B., van Eyndhoven, W.G., Nguyen, K.C.Q., Rubin, M., Hoey, T., Powers, S. and Li, J.
JOURNAL	PRC17, a novel oncogene encoding a Rab GTPase-activating protein, is amplified in prostate cancer
REFERENCE	Cancer Res. (2002) In press
AUTHORS	2 (bases 1 to 1964)
TITLE	Li, J.L., Peng, Y. and Powers, S.
JOURNAL	Direct Submission
FEATURES	Submitted (23-AUG-2002) Genomics, Tularik, Inc, 266 E Pulaski Road, Suite 1, Greenlawn, NY 11740, USA
source	Location/Qualifiers
	1. .1964

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
1..1964
/gene="PRC17"
1..1650
CDS

```

```

/gene="PRC17"
/codon_start=1
/product="Rab GTPase-activating protein PRC17"
/protein_id="AAN33117.1"
/db_xref="GI:23452665"

```

```

/translation="MDVVEVAGSWWAQEREDIIMKYEKGRAGLPEDKCPKPPRSYNN
NVDHLGIYHETELPPLTAREAKQIRREISRSKXVMDLGDMEKYSRKLIDRAYKGM
PMNIRGPMWSVLLNIEEMKLNPGRYQIMKEGKRSEHIÖRIDRVSQTLTKHIIFFR
DRYGTÖKÖREMLHLLAYEBKNPEYVQCRDLSHIAALFLYLPEBDAFMALVQLLASER
HSLÖGFHSFNGSTVÖQLÖDOQEHVAVATSPXTMGHÖDKDLGÖCSPLCGLIRLIDG
ISLGLTLRLMDVYLVEGEQALMPITRIAFVYQÖKRLTKTSRGGPMARFNRFVTWAR
DEDITVLRKLIRASMKLTKRQÖGLPPRAKEÖGSSASRPVPASRGKTKLCKGDROAPG
PPARPFRPLRWASPPRAPRSTPCPGAVREDYTPVGÖGVPSPALAÖGÖGWSRFL
ÖWNSMPRLPTLDVEGPMFRHYDPCGSCWYRAISOEDQLAPCWQAEHPABRVRSAFAA
PSTSDÖGÖTPFRARDEÖQCAPTSGPCLCGHLLESSÖFPPGF"

```

ORIGIN

Query Match	100.0%;	Score 1964;	DB 8;	Length 1964;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1964; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	1	ATGACCGTGTAGAGGTCCGGGCACTTGGTGGGCACCAAGACGAGACATCATTAAG	60
Db	1	ATGGACGTGTAGAGGTCCGGGCACTTGGTGGGCACCAAGACGAGACATCATTAAG	60
QY	61	AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA	120
Db	61	AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGGACAAAGGGCCTTAAGCCTTTTGA	120
QY	121	AGCTACAAACAACGTCGATCATTTGGGGATTGTACATGACACGAGCTGCCTCCTG	180
Db	121	AGCTACAAACAACGTCGATCATTTGGGGATTGTACATGACACGAGCTGCCTCCTG	180
QY	181	ACTGCGCGGGAAGCGGAAGCAAATTCGGCGGGAGATCAGCCGAAAGACCAAGTGGGTGAT	240
Db	181	ACTGCGCGGGAAGCGGAAGCAAATTCGGCGGGAGATCAGCCGAAAGACCAAGTGGGTGAT	240
QY	241	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAAGCTCATAGATCGAGCGTACAAG	300
Db	241	ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAAGCTCATAGATCGAGCGTACAAG	300
QY	301	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTGAACATTGAGGAATG	360
Db	301	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCACTCCTGAACATTGAGGAATG	360
QY	361	AACTTGA AAAACCCCGGAAGATACCAATCATGAAGGAAAGGCAAGAGGTCACTGAG	420
Db	361	AACTTGA AAAACCCCGGAAGATACCAATCATGAAGGAAAGGCAAGAGGTCACTGAG	420
QY	421	CACATCAGCGCATCGACCCGGGACGTAAGCGGGACATTAAAGGAACATATATTCTTCAGG	480

D	b	421	CACATCCAGCGCATCGACCCGGGACGTAAGCGGGACATTAAGGAGCATATATTCTTCAGG	480
Q	y	481	GATCGATACGGAAACCAAGCAGCGGGAACTACTCCACATCCTCCTGGCATATGAGGAGTAT	540
D	b	481	GATCGATACGGAAACCAAGCAGCGGGAACTACTCCACATCCTCCTGGCATATGAGGAGTAT	540
Q	y	541	AAACCCGAGGTGGGCTACTGCAAGGAGACCTGAGCCACATCGCCGCTTGTTCTCCTCTAT	600
D	b	541	AAACCCGAGGTGGGCTACTGCAAGGAGACCTGAGCCACATCGCCGCTTGTTCTCCTCTAT	600
Q	y	601	CTTCCTGAGGAGGATGCATTCTGGGCACTGCTGCACTGCTGCGCAGTGAGAGGCACTCC	660
D	b	601	CTTCCTGAGGAGGATGCATTCTGGGCACTGCTGCACTGCTGCGCAGTGAGAGGCACTCC	660
Q	y	661	CTGCAAGGATTTACAGCCCCAAATGGCGGAGCCGTTCCAGGGGCTCCAAAGCCAAACAGAG	720
D	b	661	CTGCAAGGATTTACAGCCCCAAATGGCGGAGCCGTTCCAGGGGCTCCAAAGCCAAACAGAG	720
Q	y	721	CATGTGTAGCCACGTCACAACCCAAAGCATTGGGGCATCAGAACAAAGATCTATGT	780
D	b	721	CATGTGTAGCCACGTCACAACCCAAAGCATTGGGGCATCAGAACAAAGATCTATGT	780
Q	y	781	GGGCAGTGTTCCCGTTAGGCTGCCTCATCCGGAATTGATTGACGGGATCTCTCGGG	840
D	b	781	GGGCAGTGTTCCCGTTAGGCTGCCTCATCCGGAATTGATTGACGGGATCTCTCGGG	840
Q	y	841	CTCACCCTGCGCCTGTGGGACGTGATCTGTGTAGAAAGGCGAAACAGGCGTTGATGCCGATA	900
D	b	841	CTCACCCTGCGCCTGTGGGACGTGATCTGTGTAGAAAGGCGAAACAGGCGTTGATGCCGATA	900
Q	y	901	ACAAAGATCGCCTTTAAGTTACAGACAAGCGCCTCACGAAGACGTCAGGTGTGGCCCG	960
D	b	901	ACAAAGATCGCCTTTAAGTTACAGACAAGCGCCTCACGAAGACGTCAGGTGTGGCCCG	960
Q	y	961	TGGGCAAGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTGCTC	1020
D	b	961	TGGGCAAGTTTTTGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTGCTC	1020
Q	y	1021	AAGCATCTTAGGGCCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCAACCCCA	1080
D	b	1021	AAGCATCTTAGGGCCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCAACCCCA	1080
Q	y	1081	GCCAAACCCGAGCAGAGGCTGTCGCGATCCAGCCTGTGCCGCTTCACTGCGCGGAAG	1140
D	b	1081	GCCAAACCCGAGCAGAGGCTGTCGCGATCCAGCCTGTGCCGCTTCACTGCGCGGAAG	1140
Q	y	1141	ACCCTCTGCAAGGGGGGACAGGCGCCCTCAAGCCCCACAGCCCGGTTCCCGCGGCC	1200
D	b	1141	ACCCTCTGCAAGGGGGGACAGGCGCCCTCAAGCCCCACAGCCCGGTTCCCGCGGCC	1200
Q	y	1201	ATTGTGTCAGCTTCCCCCGCACGGGCACTTCGTTCTTCCACACCCTGTCTGTGTGGGCT	1260
D	b	1201	ATTGTGTCAGCTTCCCCCGCACGGGCACTTCGTTCTTCCACACCCTGTCTGTGTGGGCT	1260
Q	y	1261	GTCGGGAAAGACACCTACCTGTGGGCACTCAAGGTTGCCCCAGCCCGGCTGTGCTCAG	1320
D	b	1261	GTCGGGAAAGACACCTACCTGTGGGCACTCAAGGTTGCCCCAGCCCGGCTGTGCTCAG	1320
Q	y	1321	GGAGACCTCAGGGTCTCTGGAAGATTCTGCACTGGAATCCATGCCCCGCTCCCAACG	1380
D	b	1321	GGAGACCTCAGGGTCTCTGGAAGATTCTGCACTGGAATCCATGCCCCGCTCCCAACG	1380
Q	y	1381	GACCTGGAAGTAGAGGCGCTTGTTCCGCAATTATGATTTACAGACAGAGCTGTGGGTC	1440
D	b	1381	GACCTGGAAGTAGAGGCGCTTGTTCCGCAATTATGATTTACAGACAGAGCTGTGGGTC	1440
Q	y	1441	CGTGCCATATCCAGAGGACACAGCTGGCCCTGCTGGCAGGCTGAACACCCTGCGAG	1500
D	b	1441	CGTGCCATATCCAGAGGACACAGCTGGCCCTGCTGGCAGGCTGAACACCCTGCGAG	1500
Q	y	1501	CGGTGAGATCGGCTTTCGCTGCAACCAAGCACTGATTCGACACAGGGCACCCCTTGAGA	1560

Db 1501 CCGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCCGACCAAGGGCAACCCCTTCAGA 1560
Qy 1561 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTTCGGGCTTCACTTG 1620
Db 1561 GCTAGGAGCAACAGCCGTGTGCTCCACCTCAGGGCCTTGCTTCGGGCTTCACTTG 1620
Qy 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT 1680
Db 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT 1680
Qy 1681 AATTCCCTAGGCTTACAAACCCAGCAAGCTTCGCGTCTCGTTTATTTTGGTTAAA 1740
Db 1681 AATTCCCTAGGCTTACAAACCCAGCAAGCTTCGCGTCTCGTTTATTTTGGTTAAA 1740
Qy 1741 CTTATGAATAATGTTAAGAAAGTGCAGCTCGAGAGATTCAGAGATGGAACACACC 1800
Db 1741 CTTATGAATAATGTTAAGAAAGTGCAGCTCGAGAGATTCAGAGATGGAACACACC 1800
Qy 1801 AGACCCCAAGATCACAAAAGCCAAAGCCCAAGCCCTCCCAAGCAACCCCAAGCCCAAGC 1860
Db 1801 AGACCCCAAGATCACAAAAGCCAAAGCCCAAGCCCTCCCAAGCAACCCCAAGCCCAAGC 1860
Qy 1861 CATCGTTCGAATTCCTGACGACACCGTGAGCCCTGCTTGTACTTTAACTCATGGAAG 1920
Db 1861 CATCGTTCGAATTCCTGACGACACCGTGAGCCCTGCTTGTACTTTAACTCATGGAAG 1920
Qy 1921 ATAACCTCTTACCGTTTGAATAATGTTTCTGTTGAATG 1964
Db 1921 ATAACCTCTTACCGTTTGAATAATGTTTCTGTTGAATG 1964

RESULT 3
AX775947 1964 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775947
DEFINITION Sequence 217 from Patent WO03048202.
ACCESSION AX775947
VERSION AX775947.1 GI:32693665

KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homindae; Homo.

REFERENCE
1 Matsuda, A. and Muramatsu, S.
AUTHORS NF-kB activating gene
TITLE Patent: WO 03048202-A 217 12-JUN-2003;
JOURNAL Asahi Kasei Kabushiki Kaisha (JP)

FEATURES
source
1. 1964
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 1650
/note="unnamed protein product"

/codon_start=1
/protein_id="CAB11673.1"
/db_xref="GI:32693666"
/translation="MDVVEVAGSWMWAOEREDIMKYEGHRAGLPEDKPKPFRSYNN
NVDHLGI VHETELPLTAREAKQIREISRKSYDMLGDWEKYKSRKLI DRAYKGM
PMNIRGPMWSVLNIEEMKLNKPRYQIMKEKRSSEHIIQIRIDRVSGTLRKHIFR
DRYGTQKRELHLILAYEYNPEVGYCRDLSHIALFLYLPEEDAFWALVQLASER
HSLGFHSPNGGTGLODOQEHVATISQPKTMGHODKDLGQCSPLGCLIRILIDG
ISLGLTRLMDVYLVEGEQALMPITRIAFKVQQRKLTXTSRGPMARFCNRFVDTWAR
DEDTVLKHLRASMKLTKRQGLQRPAPKEQGSASRPVPSRGSKTLCKGDROAPG
PPARFPRPIWSASPPRAPRSSTPCPGGAVREDTYVPVGTQVPSPALAQGQFQGSWRF
QMSMPRLPTDL DVEGPMFRHYDFRQSCWVRATISQEDQLAPCMQAEHPAERVRSAPFA
PSTDSQGT PFRARDEQPCAPTSGPCLCGLHBS SQFP PGF"

ORIGIN
Query Match 99.9%; Score 1962.4; DB 6; Length 1964;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1963; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGACGTGTAGAGTTCGCGGCGAGTTGTGGGCAAAAGCGAGAGACATCATTTATG 60
Db 1 ATGACGTGTAGAGTTCGCGGCGAGTTGTGGGCAAAAGCGAGAGACATCATTTATG 60
Qy 61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGGCAAAAGGGCCTTAAGCTTTTGA 120
Db 61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGGCAAAAGGGCCTTAAGCTTTTGA 120
Qy 121 AGCTACAAACAACAGCTGCATCATTTGGGATTTGTACATGACGAGCTGCTCTCTG 180
Db 121 AGCTACAAACAACAGCTGCATCATTTGGGATTTGTACATGACGAGCTGCTCTCTG 180
Qy 181 ACTGCGGGGAGGCGGAAGCAAAATTCGGCGGGAGATCAGCCGAAAAGCAAGTGGTGAT 240
Db 181 ACTGCGGGGAGGCGGAAGCAAAATTCGGCGGGAGATCAGCCGAAAAGCAAGTGGTGAT 240
Qy 241 ATGCTGGGAGACTGGAGAAATACAAAAGCAGAGAAAGCTCATAGATCGAGCTACAG 300
Db 241 ATGCTGGGAGACTGGAGAAATACAAAAGCAGAGAAAGCTCATAGATCGAGCTACAG 300
Qy 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAAACATTGAGAAATG 360
Db 301 GGAATGCCCATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAAACATTGAGAAATG 360
Qy 361 AAGTTGAAAAAAGCCCGGAGATACAGATCATGAAGAGAAAGGCAAGAGTCATCTGAG 420
Db 361 AAGTTGAAAAAAGCCCGGAGATACAGATCATGAAGAGAAAGGCAAGAGTCATCTGAG 420
Qy 421 CACATCCAGCGCATCGACCGGGACGTAAGCGGGACATTAAGGAAGCATATTTCTCAGG 480
Db 421 CACATCCAGCGCATCGACCGGGACGTAAGCGGGACATTAAGGAAGCATATTTCTCAGG 480
Qy 481 GATCGATACGGAACCAAGCAGCGGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT 540
Db 481 GATCGATACGGAACCAAGCAGCGGGAACTACTCCACATCTCTCTGGCATATGAGAGTAT 540
Qy 541 AACCCGAGGTGGGCTACTGAGGGAACCTGAGCCATTCGCCGCTTGTCTCTCTAT 600
Db 541 AACCCGAGGTGGGCTACTGAGGGAACCTGAGCCATTCGCCGCTTGTCTCTCTAT 600
Qy 601 CTTCTGAGGAGATGATCTGGGCACTGTGCTGCTGGCCAGTGAAGGCACTCC 660
Db 601 CTTCTGAGGAGATGATCTGGGCACTGTGCTGCTGGCCAGTGAAGGCACTCC 660
Qy 661 CTGCAAGGATTTCAACAGCCCAAATGCGGGGACCCGTCCAAAGCAACAGAG 720
Db 661 CTGCAAGGATTTCAACAGCCCAAATGCGGGGACCCGTCCAAAGCAACAGAG 720
Qy 721 CATGTGTAGCCACGTCACAAACCCAAAGCCATGGGGCATCAGACAAAGAAATCTATGT 780
Db 721 CATGTGTAGCCACGTCACAAACCCAAAGCCATGGGGCATCAGACAAAGAAATCTATGT 780
Qy 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTTGATTCAGGGATCTCTCGGG 840
Db 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATTTGATTCAGGGATCTCTCGGG 840
Qy 841 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGCGTTGATGCGGATA 900
Db 841 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGCGTTGATGCGGATA 900
Qy 901 ACAAGATCGCCTTTAAGTTCAAGCAGAGCGCCTCAAGAGAGCTCCAGGTGCGCCG 960
Db 901 ACAAGATCGCCTTTAAGTTCAAGCAGAGCGCCTCAAGAGAGCTCCAGGTGCGCCG 960
Qy 961 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCCAAGGATGAGGACACTGTGCTC 1020
Db 961 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCCAAGGATGAGGACACTGTGCTC 1020
Qy 1021 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1021 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCACCCCA 1080

QY	1081	GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCGCTGTGCGGCTTCACTGCGGGAAG	1140
Db	1081	GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCGCTGTGCGGCTTCACTGCGGGAAG	1140
QY	1141	ACCCTCTGCAAGGGGACAGGACAGGCGCCCTCCAGGCCCAACGCGGTTCCCGCGGCC	1200
Db	1141	ACCCTCTGCAAGGGGACAGGACAGGCGCCCTCCAGGCCCAACGCGGTTCCCGCGGCC	1200
QY	1201	ATTGTGTCAGCTTCCCCGCCACGCGGCACCTCGTTCTTCCACACCTGTCTGTGGGCT	1260
Db	1201	ATTGTGTCAGCTTCCCCGCCACGCGGCACCTCGTTCTTCCACACCTGTCTGTGGGCT	1260
QY	1261	GTCCGGGAAGACACCTTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG	1320
Db	1261	GTCCGGGAAGACACCTTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG	1320
QY	1321	GGAGGACCTCAGGGTCTGAGAGATTCTGACGTGAATCTCCATGCCCCGCTCCCAAG	1380
Db	1321	GGAGGACCTCAGGGTCTGAGAGATTCTGACGTGAATCTCCATGCCCCGCTCCCAAG	1380
QY	1381	GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTAGACAGAGCTGTGGTC	1440
Db	1381	GACCTGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTAGACAGAGCTGTGGTC	1440
QY	1441	CGTGCCATATCCCAAGAGAACAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTCGAG	1500
Db	1441	CGTGCCATATCCCAAGAGAACAGCTGGCCCCCTGCTGGCAGGCTGAACACCCCTCGAG	1500
QY	1501	CGGGTGAATCGGCTTTCGCTGCACCCAGCAGCTGAATTCGACACAGGGCACCCCTTCAGA	1560
Db	1501	CGGGTGAATCGGCTTTCGCTGCACCCAGCAGCTGAATTCGACACAGGGCACCCCTTCAGA	1560
QY	1561	GCTAGGACGAACAGCCGTGTGCTCCCACTCAGGGCCCTGCTGTGCGGCTCCAATTG	1620
Db	1561	GCTAGGACGAACAGCCGTGTGCTCCCACTCAGGGCCCTGCTGTGCGGCTCCAATTG	1620
QY	1621	GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT	1680
Db	1621	GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT	1680
QY	1681	AATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGCGTCTCGTTTATTTTGTAA	1740
Db	1681	AATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGCGTCTCGTTTATTTTGTAA	1740
QY	1741	CTTATGAATAATGTATTAAAGAAAGATGCAGCTCGAAGAGATTCAAGATGGAACACACC	1800
Db	1741	CTTATGAATAATGTATTAAAGAAAGATGCAGCTCGAAGAGATTCAAGATGGAACACACC	1800
QY	1801	AGACCCAGATCACAAAGCCCAACCATGCCAGCCCCCTCCAGCACCCCAAGCCCAAGAC	1860
Db	1801	AGACCCAGATCACAAAGCCCAACCATGCCAGCCCCCTCCAGCACCCCAAGCCCAAGAC	1860
QY	1861	CATCGTTCTGAATTCTGACGACACCGTGAGCCTTGTACTTTAACTCATGAAGG	1920
Db	1861	CATCGTTCTGAATTCTGACGACACCGTGAGCCTTGTACTTTAACTCATGAAGG	1920
QY	1921	ATAACTACCTTCAAGTTTGAATAAATGTTTCTGTGAATG	1964
Db	1921	ATAACTACCTTCAAGTTTGAATAAATGTTTCTGTGAATG	1964
RESULT 4			
LOCUS	BC071680	2090 bp	mRNA linear PRI 08-MAR-2005
DEFINITION	Homo sapiens TBCL domain family, member 3, mRNA (cDNA clone		
ACCESSION	MGC:87891 IMAGE:5165385), complete cds.		
VERSION	BC071680		
KEYWORDS	BC071680.1 GI:47939522		
SOURCE	MGC.		
ORGANISM	Homo sapiens (human)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		

REFERENCE	1 (bases 1 to 2090)
AUTHORS	Strauberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altshul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaly SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodriguez S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smalius DE, Schnerch A, Schein JE, Jones SJ and Marra MA.
CONSRM	Mammalian Gene Collection Program Team
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE	2 (bases 1 to 2090)
AUTHORS	Director MGC Project.
TITLE	Direct Submission
JOURNAL	Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Location/Qualifiers
source	1..2090
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="MGC:87891 IMAGE:5165385"
	/tissue_type="Brain, adult medulla"
	/clone_id="NIH_MGC_119"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"
gene	1..2090
	/gene="TBC1D3"
	/note="synonyms: PRC17, TRE17"
	/db_xref="GeneID:84218"
	/db_xref="MIM:607741"
	101..1750
	/gene="TBC1D3"
	/codon_start=1
	/product="TBC1 domain family, member 3"
	/protein_id="AAH71680.1"
	/db_xref="GI:47939523"
	/db_xref="GeneID:84218"
	/db_xref="MIM:607741"
	/translation="MDVVEVAGSWWAQERDITMKYKGRAGLPEDKPKPFRSYNNVNDHGIIVHETELPPLTAREBAKQIRREISRKSKWVMDLGWEKYKSRKLDRAVYKGMFNNIRGPMWSVLNIEEMKLKNPGRYQIMKEKGRSSSEHIQRIDRDISGTLRKHMFFR
CDS	

DRYGTKORELIHILLAYEENPEVGYCRDLSHIALFLYLPEEDAFWALVOLLASER
HSLQFHSNPGSTVQGLQDQOEHVAVTSQPKTMGHQDKDLQCSPLGLIRLIDG
ISLGLTLRLMDVLYVEGEOLMPTIRIAFKQOKRLTKISRCPWAFCNRFVDTWAR
DEDTVLKLRLASMKULTRKQDLPFPAKREQSSASRPVPASRGKTLCKDRQAPR
PPARFPRPWSASPPARPRSSTPCPGAVREDTYPVGTOGVSPLAQGGPOGSMRFL
QWNSMPRLPTDLIDVEGPMFRHYDFRQSCVWRAISQEDOLAPCWOAEHPAERVRSAPAA
PSTSDQGTPPRARDEQQCAPTSGPCLGLHLESSQFP PGF"

ORIGIN
Query Match 99.4%; Score 1952.8; DB 8; Length 2090;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1957; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY	1	ATGACGTGTAGAGGTGCGGGCAGTTGTGGGCAAGA	CGAGAGACATCATTTATG	60
Db	101	ATGACGTGTAGAGGTGCGGGCAGTTGTGGGCAAGA	CGAGAGACATCATTTATG	160
QY	61	AAATACGAAAAGGACACCGAGCTGGGCTGCCAGA	GACAAGGGGCTAAGCCTTTTGA	120
Db	161	AAATACGAAAAGGACACCGAGCTGGGCTGCCAGA	GACAAGGGGCTAAGCCTTTTGA	220
QY	121	AGCTACAACAACAACGTTCATCTTTGGGATTTGT	ACATGAGACGGAGCTGCTCTG	180
Db	221	AGCTACAACAACAACGTTCATCTTTGGGATTTGT	ACATGAGACGGAGCTGCTCTG	280
QY	181	ACTGCGCGGAGCGGAAGCAAATTCGGCGGAGAT	CAGCCGAAAAGCAAGTGGTGAT	240
Db	281	ACTGCGCGGAGCGGAAGCAAATTCGGCGGAGAT	CAGCCGAAAAGCAAGTGGTGAT	340
QY	241	ATGCTGGGAGACTGGAGAAATACAAAAGCAGCA	GAAGAACTCATAGACGCTACAG	300
Db	341	ATGCTGGGAGACTGGAGAAATACAAAAGCAGCA	GAAGAACTCATAGACGCTACAG	400
QY	301	GGAATGCCCATGAACATCCGGGGCCCGATGTGT	CACTCCTGAAACATGAGAAATG	360
Db	401	GGAATGCCCATGAACATCCGGGGCCCGATGTGT	CACTCCTGAAACATGAGAAATG	460
QY	361	AAATTGAAAAAACC	CGGAAGATACAGATCATGAAAGAAAGGCAAG	420
Db	461	AAATTGAAAAAACC	CGGAAGATACAGATCATGAAAGAAAGGCAAG	520
QY	421	CACATCCAGCGCATCGACCGGGA	CGTTAAGCGGACATTTAAGAAAGCATATTTCT	480
Db	521	CACATCCAGCGCATCGACCGGGA	CGTTAAGCGGACATTTAAGAAAGCATATTTCT	580
QY	481	GATCGATACGGAACCAAGCAGCGGGA	ACTTACTCCATCTCTTGCCATATGAGGATAT	540
Db	581	GATCGATACGGAACCAAGCAGCGGGA	ACTTACTCCATCTCTTGCCATATGAGGATAT	640
QY	541	AA	CCGGAGGTGGGCTACTGCAAGGAACTTGAGCCACATCGCCGCTTGTCTCTCTAT	600
Db	641	AA	CCGGAGGTGGGCTACTGCAAGGAACTTGAGCCACATCGCCGCTTGTCTCTCTAT	700
QY	601	CTTCTGAGAGAGATGATCTGGGCACTGTGCA	GCTGCTGGCCAGTGAGGCACTCC	660
Db	701	CTTCTGAGAGAGATGATCTGGGCACTGTGCA	GCTGCTGGCCAGTGAGGCACTCC	760
QY	661	CTGCAAGGATTTACAGGCCCAAATGGCGGGA	ACCGTCCAGGGGCTCCAAAGCCAAGAG	720
Db	761	CTGCAAGGATTTACAGGCCCAAATGGCGGGA	ACCGTCCAGGGGCTCCAAAGCCAAGAG	820
QY	721	CATGTGTAGCCACGTACAAACCAAGACCAT	TGGGCAATGAGACAAAGAAAGATCTATGT	780
Db	821	CATGTGTAGCCACGTACAAACCAAGACCAT	TGGGCAATGAGACAAAGAAAGATCTATGT	880
QY	781	GGGCAGTGTCCCGTTAGGCTGCTCATCCGAT	ATTTGATGACGGATCTCTCGGG	840
Db	881	GGGCAGTGTCCCGTTAGGCTGCTCATCCGAT	ATTTGATGACGGATCTCTCGGG	940
QY	841	CTCACCTGCGCCTGTGGGACGTGTATCTGT	AGAAAGGCAACAGCGTTGATGCCGATA	900
Db	941	CTCACCTGCGCCTGTGTGGACGTGTATCTGT	AGAAAGGCAACAGCGTTGATGCCGATA	1000

QY	901	ACAAGATCGCCTTTAAAGTTACAGCAGAAAGC	CCCTCACGAAAGCTCCAGGTGTGGCCG	960
Db	1001	ACAAGATCGCCTTTAAAGTTACAGCAGAAAGC	CCCTCACGAAAGCTCCAGGTGTGGCCG	1060
QY	961	TGGGCACTTTTGGCAACGGTTCGTTGATAC	TGGGCAAGGATGAGAACACTGTGCTC	1020
Db	1061	TGGGCACTTTTGGCAACGGTTCGTTGATAC	TGGGCAAGGATGAGAACACTGTGCTC	1120

QY	1021	AAGCATCTTAGGCTCTATGAAAGAACTAACAA	AGAAAGCAGGGGCACTGCCACCCCA	1080
Db	1121	AAGCATCTTAGGCTCTATGAAAGAACTAACAA	AGAAAGCAGGGGCACTGCCACCCCA	1180
QY	1081	GCCAAACCCGAGCAAGGTCGTCCGATCCAG	GCCTGTGCGGCTTCACTGCGGGAAG	1140
Db	1181	GCCAAACCCGAGCAAGGTCGTCCGATCCAG	GCCTGTGCGGCTTCACTGCGGGAAG	1240
QY	1141	ACCCTCTGCAAGGGGGACAGGCAAGGCCCT	CCAGGCCCAACAGCCCGTTCCCGGGCC	1200
Db	1241	ACCCTCTGCAAGGGGGACAGGCAAGGCCCT	CCAGGCCCAACAGCCCGTTCCCGGGCC	1300
QY	1201	ATTGTGACGTTCCCCGCCACGGGCACTG	TTCTTCCACACCCCTGTGTTGGGCT	1260
Db	1301	ATTGTGACGTTCCCCGCCACGGGCACTG	TTCTTCCACACCCCTGTGTTGGGCT	1360
QY	1261	GTC	CGGGAAGACACTACCTGTGGGCACTCAGG	1320
Db	1361	GTC	CGGGAAGACACTACCTGTGGGCACTCAGG	1420
QY	1321	GGAGACCTCAGGGTCTTGAGATTCTGCA	GTGAACTCCATGCCCGCTCCCAACG	1380
Db	1421	GGAGACCTCAGGGTCTTGAGATTCTGCA	GTGAACTCCATGCCCGCTCCCAACG	1480
QY	1381	GACCTGACGTAGAGGGCCCTTGCTCCG	CAATATGATTTACAGACAGAGCTGGGTC	1440
Db	1481	GACCTGACGTAGAGGGCCCTTGCTCCG	CAATATGATTTACAGACAGAGCTGGGTC	1540
QY	1441	CGTGCCATATCCCAAGAGAACAGCTG	GGCCCCCTGCTGGCAGGCTGAACAC	1500
Db	1541	CGTGCCATATCCCAAGAGAACAGCTG	GGCCCCCTGCTGGCAGGCTGAACAC	1600
QY	1501	CGG	TGAGATCGGCTTTGCTGCAACCCAGCA	1560
Db	1601	CGG	TGAGATCGGCTTTGCTGCAACCCAGCA	1660
QY	1561	GCTAGGAGCAACAGCCG	TGTGCTCCCACTCAGGGCCTTGCCCTG	1620
Db	1661	GCTAGGAGCAACAGCCG	TGTGCTCCCACTCAGGGCCTTGCCCTG	1720
QY	1621	GAAAGTTCTCAGTTCCCTCCAGGCTT	TAGAAGCATCTGGCCAGGCTCATGG	1680
Db	1721	GAAAGTTCTCAGTTCCCTCCAGGCTT	TAGAAGCATCTGGCCAGGCTCATGG	1780
QY	1681	AA	TTTCCCTAGGCTTAACAACCAAGCAAG	1740
Db	1781	AA	TTTCCCTAGGCTTAACAACCAAGCAAG	1840
QY	1741	CTTATGAAATGTATTAAAGAAAGAGT	GCAGCTGAGAGAGATTACAGATGGA	1800
Db	1841	CTTATGAAATGTATTAAAGAAAGAGT	GCAGCTGAGAGAGATTACAGATGGA	1900
QY	1801	AGA	CCCCAGATCACAAAGCCAACATG	1860
Db	1901	AGA	CCCCAGATCACAAAGCCAACATG	1960
QY	1861	CATGTTCTGAATCTGAGACA	CCGTCGCTTTGTACTTTAACTCAT	1920
Db	1961	CATGTTCTGAATCTGAGACA	CCGTCGCTTTGTACTTTAACTCAT	2020
QY	1921	ATA	CTACCTTCACGTTTGAATAATATG	1984
Db	2021	ATA	CTACCTTCATGTTTTGAATAATATG	2064

```
RESULT 5
AX086847
LOCUS AX086847 2072 bp DNA linear PAT 09-MAR-2001
DEFINITION Sequence 799 from Patent WO0112659.
ACCESSION AX086847
VERSION AX086847.1 GI:13276047
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1
AUTHORS Wiemann,S.
TITLE Human dna sequences
JOURNAL Patent: WO 0112659-A 799 22-FEB-2001;
German Human Genome Project (DE)
FEATURES
source
1.2072
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.2%; Score 1948; DB 6; Length 2072;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1954; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 ATGACGTGTAGAGGTGCGGGGCACTTGGTGCGACACAAGCGAGACATCATTTATG 60
DB 99 ATGACGTGTAGAGGTGCGGGGCACTTGGTGCGACACAAGCGAGACATCATTTATG 158
QY 61 AAATACGAAGGAGACACCGAGCTGGGCTGCCAGAGACAAGGGCCTTAAGCCTTTTGA 120
DB 159 AAATACGAAGGAGACACCGAGCTGGGCTGCCAGAGACAAGGGCCTTAAGCCTTTTGA 218
QY 121 AGCTACAACAACAACGTGCATCTTGGGGATTGTACATGAGACGAGCTGCCCTCTG 180
DB 219 AGCTACAACAACAACGTGCATCTTGGGGATTGTACATGAGACGAGCTGCCCTCTG 278
QY 181 ACTGCGGGGAGCGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 240
DB 279 ACTGCGGGGAGCGGAGCAAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGTGAT 338
QY 241 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAAAAGCTCATAGATCGAGCGTACAAG 300
DB 339 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAAAAGCTCATAGATCGAGCGTACAAG 398
QY 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 360
DB 399 GGAATGCCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 458
QY 361 AAGTTGAAAAAACC CGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCACTGAG 420
DB 459 AAGTTGAAAAAACC CGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCACTGAG 518
QY 421 CACATCCAGCGCATCGACCGGAGACGTAAAGCGGACATTAAGGAACATATATTCTTCAAG 480
DB 519 CACATCCAGCGCATCGACCGGAGACGTAAAGCGGACATTAAGGAACATATATTCTTCAAG 578
QY 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 540
DB 579 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 638
QY 541 AACC CGAGGTGGCTACTGCAAGGACCTGAGCCACATCGCGCTTGTCTCTCTAT 600
DB 639 AACC CGAGGTGGCTACTGCAAGGACCTGAGCCACATCGCGCTTGTCTCTCTAT 698
QY 601 CTTCTGAGGAGATGATCTTGCGCACTGGTGACGTGCTGGCCAGTGAGAGCACTCC 660
DB 699 CTTCTGAGGAGATGATCTTGCGCACTGGTGACGTGCTGGCCAGTGAGAGCACTCC 758
QY 661 CTGAGGAGATTTCACAGCCAAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
```

```
DB 759 CTGACGGATTTTCACAGCCCAATGGCGGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 818
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 780
DB 819 CATGTGTAGCCACGTCAACAACCAAGACCATGGGGCATCAGGACAAGAAAGATCTATGT 878
QY 781 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTTGAACGGGATCTCTCGGG 840
DB 879 GGGCAGTGTCCCCGTTAGGCTGCTCATCCGATATTTGAACGGGATCTCTCGGG 938
QY 841 CTCACCTGCGCTGTGAGACGTGTATCTGTAGAAAGCGAAGCGGCTGATGCCGATA 900
DB 939 CTCACCTGCGCTGTGAGACGTGTATCTGTAGAAAGCGAAGCGGCTGATGCCGATA 998
QY 901 ACAGAATCCCTTTAAGGTTACAGAGAAGCGCCTCACGAAGACGTCAGGTGCGCCG 960
DB 999 ACAGAATCCCTTTAAGGTTACAGAGAAGCGCCTCACGAAGACGTCAGGTGCGCCG 1058
QY 961 TGGGACGTTTTTGCACACCGGTTGCTGATACCTGGGCAAGGATGAGACACTGTGCTC 1020
DB 1059 TGGGACGTTTTTGCACACCGGTTGCTGATACCTGGGCAAGGATGAGACACTGTGCTC 1118
QY 1021 AAGCATCTTAGGCGCTCTATAGAAATACTAACAAAGAAAGAGGGGACCTGCAACCCCA 1080
DB 1119 AAGCATCTTAGGCGCTCTATAGAAATACTAACAAAGAAAGAGGGGACCTGCAACCCCA 1178
QY 1081 GCCAAACCGGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTACGTGGCGGAAG 1140
DB 1179 GCCAAACCGGAGCAAGGTCGTGGCATCCAGGCTGTGCGGCTTACGTGGCGGAAG 1238
QY 1141 ACCCTCTGCAAGGGGAGACAGGAGCCCTCCAGGCGCCACAGCCCGGTTCCCGCGCC 1200
DB 1239 ACCCTCTGCAAGGGGAGACAGGAGCCCTCCAGGCGCCACAGCCCGGTTCCCGCGCC 1298
QY 1201 ATTTGTCAGCTTCCCCGCCACGGGACCTCGTTCTTCCACACCCTGTCTGTGGGGCT 1260
DB 1299 ATTTGTCAGCTTCCCCGCCACGGGACCTCGTTCTTCCACACCCTGTCTGTGGGGCT 1358
QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGCCCCAGCCCGGCTGCTCAG 1320
DB 1359 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGCCCCAGCCCGGCTGCTCAG 1418
QY 1321 GGAAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCCATGCCCCGCTCCCAAG 1380
DB 1419 GGAAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAATCCATGCCCCGCTCCCAAG 1478
QY 1381 GACCTGACGTAGAGGGCCCTTGTCCGCCATTATGATTTCAACAAGAGCTGTGGTTC 1440
DB 1479 GACCTGACGTAGAGGGCCCTTGTCCGCCATTATGATTTCAACAAGAGCTGTGGTTC 1538
QY 1441 CGTGCCATATCCAGAGGACAGCTGGCCCCCTGTGGCAAGGCTGAACACCTGCGGAG 1500
DB 1539 CGTGCCATATCCAGAGGACAGCTGGCCCCCTGTGGCAAGGCTGAACACCTGCGGAG 1598
QY 1501 CGGCTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCGACAGGGCACCCCTTCAGA 1560
DB 1599 CGGCTGAGATCGGCTTTCGCTGCAACCCAGCACTGATTCGACAGGGCACCCCTTCAGA 1658
QY 1561 GCTAGGAGCAACAGCGGTGTCTCCACCTCAGGGCTTGCTCTGCGGCTCACTTG 1620
DB 1659 GCTAGGAGCAACAGCGGTGTCTCCACCTCAGGGCTTGCTCTGCGGCTCACTTG 1718
QY 1621 GAAATTTCTCAGTTCCCTCAAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGCTGAT 1680
DB 1719 GAAATTTCTCAGTTCCCTCAAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGCTGAT 1778
QY 1681 AATTTCCCTAGGCTTAACAACCAAGCAAGCTTCCGCTCTCTTTATTTTGGTTAAA 1740
DB 1779 AATTTCCCTAGGCTTAACAACCAAGCAAGCTTCCGCTCTCTTTATTTTGGTTAAA 1838
QY 1741 CTTATGAAAAATGTATTAAGAAAGAGTGACGCTCGAGAGAGATTCAAGAGTGAACACACC 1800
```


Db 1839 CTTATGAAATGTATTAGAAAGAGTGCAGCTCGAGAGATTCAAGATGGAACACACC 1898
Qy 1801 AGACCCCATGATCACAAGCCCAACCATGCCCCCCTCCAGACACCCCGAGCCCGACGAC 1860
Db 1899 AGACCCCATGATCACAAGCCCAACCATGCCCCCCTCCAGACACCCCGAGCCCGACGAC 1958
Qy 1861 CATCGTTGTAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTAACTCATGGAAG 1920
Db 1959 CATCGTTGTAATTTCTGACGACACCGTGAGCCTGCTTTGTACTTTCAACTCATGGAAG 2018
Qy 1921 ATAACTACTTCAAGTTTGAATAATGTTTCTGTGAAATG 1964
Db 2019 ATAACTACTTCAAGTTTGAATAATGTTTCTGTGAAATG 2062

RESULT 6
LOCUS BC075809 2166 bp mRNA linear PRI 17-MAY-2005
DEFINITION Homo sapiens TBC1 domain family, member 3B, mRNA (cDNA clone
ACCESSION BC075809 MGC:87892 IMAGE:5170064), complete cds.
VERSION BC075809.1 GI:49902676
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE 1 (bases 1 to 2166)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Utsdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,B., Kettelman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
CONSRM Mammalian Gene Collection Program Team
TITLE Generation and initial analysis of more than 15,000 full-length
JOURNAL human and mouse cDNA sequences
PUBMED Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
REFERENCES 12477932
2 (bases 1 to 2166)
CONSRM NIH MGC Project
TITLE Direct Submission
AUTHORS Submitted (06-JUL-2004) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-ehgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov

Series: IRAK Plate: 166 Row: h Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 14149984.
FEATURES
source location/Qualifiers
1..2166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87892 IMAGE:5170064"
/tissue_type="Brain, adult medulla"
/clone_1ib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2166

gene
/gene="TBC1D3B"
/note="synonym: PRC17"
/db_xref="GeneID:414059"
147..1796
/gene="TBC1D3B"
/codon_start=1
/product="TBC1 domain family, member 3B"
/protein_id="AAH75809.1"
/db_xref="GI:49902677"
/db_xref="GeneID:414059"
/translation="MDVEEVAQWEREDIMKYEGHRAGLEDPKPKPRSYNN
NYDLGI VHETELPPLTAREAKQIRREISRSKXVNDMLGDWEKYSRKLIDRAYKGM
PMNIRGPMWSVLTLEEMKLNKPGRYQIMKEKGRSSEHIQRIDRDISGTLRHRMFFR
DRYGTQKRELHILILAYEENPEVGYCRDLISHIALFLYLPEEAFWALVQLASER
HSLQGFSPNGFTVQGLQDQOEHVATVSQTMGHODKDLGCGSPGLCLIRILIDR
ISLGLTLRLMDVYLVEGEQALMPITRIAEKVQKRLTSTRCGPWAFNCFVDTWAR
DDDTVLKHLRASMKLTKRKGDLPPPAKEOGSSASRPVPASRGKTLCKGDRAPPG
PPARFPRIWSPAPRASSTPCPGAVREDTYPVGTQVPSPALAOGGPGQSWRFL
QNSMPRLPTDLDEGFWERHYDFRQSCWTVRAISOEDOLAPCWQAEHPAERVRSFAFA
PDTSDQGTFFRADEQQCAPTSGPCLGLHLESSQFPFPGF"

ORIGIN

Query Match 98.6%; Score 1936.4; DB 8; Length 2166;
Best local Similarity 99.2%; Pred. No. 0;
Matches 1946; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 1 ATGAGCGTGTAGAGGTCCGGGCGAGTTGGTGGCACAGAAGCGAGACATCATTTATG 60
Db 147 ATGAGCGTGTAGAGGTCCGGGCGAGTTGGTGGCACAGAAGCGAGACATCATTTATG 206
Qy 61 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 120
Db 207 AAATACGAAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 266
Qy 121 AGCTACACAACAACGTGCATTTGGGATTTGTACATGAGACGAGCTGCTCTCTG 180
Db 267 AGCTACACAACAACGTGCATTTGGGATTTGTACATGAGACGAGCTGCTCTCTG 326
Qy 181 ACTGCGCGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT 240
Db 327 ACTGCGCGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAGAAGCAAGTGGTGAT 386
Qy 241 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGCTACAAG 300
Db 387 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGCAGAAAGCTCATAGATCGAGCTACAAG 446
Qy 301 GGAATGCCATGAACATCCGGGGCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATG 360
Db 447 GGAATGCCATGAACATCCGGGGCCGATGTGTGTCAGTCTCTGAACATTGAGGAAATG 506
Qy 361 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCTTGAG 420
Db 507 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCTTGAG 566
Qy 421 CACATCAGCGCATGACCGGGACGTAAGCGGACATTAAGAAAGCATATATCTTCAGG 480
Db 567 CACATCAGCGCATGACCGGGACATTAAGCGGACATTAAGAAAGCATATATCTTCAGG 626
Qy 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCAATCTCTGCGATATGAGAGTAT 540

Db 627 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 686
Qy 541 AACCCGAGGTGGGCTACTGACAGGAACTGAGCCACATCGCCGCTTGTTCTCTCTAT 600
Db 687 AACCCGAGGTGGGCTACTGACAGGAACTGAGCCACATCGCCGCTTGTTCTCTCTAT 746
Qy 601 CTTCTGAGGAGATGCAATCTGGGCACTGGTGACGTGCTGGCCAGTGAAGCACTCC 660
Db 747 CTTCTGAGGAGATGCAATCTGGGCACTGGTGACGTGCTGGCCAGTGAAGCACTCC 806
Qy 661 CTGCAAGGATTTACAGCCCAAATGGCGGAGCCGTCAGGGGCTCCAAGACCAAGAG 720
Db 807 CTGCAAGGATTTACAGCCCAAATGGCGGAGCCGTCAGGGGCTCCAAGACCAAGAG 866
Qy 721 CATGTGGTAGCCACGTCACAAACCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 780
Db 867 CATGTGGTAGCCACGTCACAAATCCAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 926
Qy 781 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGGAATATGATGACGGGATCTCTCGGG 840
Db 927 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGGAATATGATGACCGGATCTCTCGGG 986
Qy 841 CTCACCTGCGCGCTGTGGAGCGTGTATCTGTAGAAGCGAAACAGGCGTGTATGCCGATA 900
Db 987 CTCACCTGCGCGCTGTGGAGCGTGTATCTGTAGAAGCGAAACAGGCGTGTATGCCGATA 1046
Qy 901 ACAAGATCGCCTTTAAGGTTACAGAGAGCGCTTCAGAAAGCGTCCAGGTGTGGCCCG 960
Db 1047 ACAAGATCGCCTTTAAGGTTACAGAGAGCGCTTCAGAAAGCGTCCAGGTGTGGCCCA 1106
Qy 961 TGGGCAAGTTTTTGCAACCGGTTGTTGATACCTGGGCCAGGATGAGGACATGTGCTC 1020
Db 1107 TGGGCAAGTTTTTGCAACCGGTTGTTGATACCTGGGCCAGGATGATGACATGTGCTC 1166
Qy 1021 AAGCATTTAGGGCCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCCCCA 1080
Db 1167 AAGCATTTAGGGCCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCCCCA 1226
Qy 1081 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCGGCTTCACGTGGCGGAAAG 1140
Db 1227 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCGGCTTCACGTGGCGGAAAG 1286
Qy 1141 ACCCTCTGCAAGGGGAGCAGGAGCCCTCCAGGCCCAACAGCCCGGTTCCCGCGGCC 1200
Db 1287 ACCCTCTGCAAGGGGAGCAGGAGCCCTCCAGGCCCAACAGCCCGGTTCCCGCGGCC 1346
Qy 1201 ATTGGTCAAGTTCCTCCGCGCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT 1260
Db 1347 ATTGGTCAAGTTCCTCCGCGCACGGGCACTCGTTCTTCCACACCCTGTCTGTGGGGCT 1406
Qy 1261 GTCGGGGAAGACACCTTACCTGTGGGCACTCAGGGTGTGCCAAGCCCGGCTGGCTCAG 1320
Db 1407 GTCGGGGAAGACACCTTACCTGTGGGCACTCAGGGTGTGCCAAGCCCGGCTGGCTCAG 1466
Qy 1321 GGAAGACCTCAGGGTTCCTGGAAGATTCTGCAGTGGAATCCATGCCCCGCTCCCAAG 1380
Db 1467 GGAAGACCTCAGGGTTCCTGGAAGATTCTGCAGTGGAATCCATGCCCCGCTCCCAAG 1526
Qy 1381 GACCTGGAAGTAGAGGCGCTGTGCTCCGCAATATGATTTAGACAGAGCTGTGGGTC 1440
Db 1527 GACCTGGAAGTAGAGGCGCTGTGCTCCGCAATATGATTTAGACAGAGCTGTGGGTC 1586
Qy 1441 CGTGCCATATCCAGAGGACCAAGCTGGCCCTGCTGGCAGGCTGAACCCCTGGCAG 1500
Db 1587 CGTGCCATATCCAGAGGACCAAGCTGGCCCTGCTGGCAGGCTGAACCCCTGGCAG 1646
Qy 1501 CGGCTGAGATCGGCTTTCGCTGCAACCCAGCATGATTCGACCAAGGGCAACCCCTTACA 1560
Db 1647 CGGCTGAGATCGGCTTTCGCTGCAACCCGGCATGATTCGACCAAGGGCAACCCCTTACA 1706
Qy 1561 GCTAAGGACGAAGCCGCTGTCTCCCACTCAGGGCCTTGCTGTGGGCTTCACTTG 1620

Db 1707 GCTAAGGACGAACAGCAGTGTGCTCCCACTCAGGGCCTTGCTGTGGGCTCCACTTG 1766
Qy 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
Db 1767 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGAT 1826
Qy 1681 AATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCGTCTCGTTTATTTTGGTTAA 1740
Db 1827 AATTTCCCTAGGCTTAACAACCCCAAGCAAGCTTCGCGTCTCGTTTATTTTGGTTAA 1886
Qy 1741 CTTATGAAAATGTATTAGAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
Db 1887 CTTATGAAATGTATTAGAAGAGTGCAGCTTGAAGACATTCAAGATGGAACACACC 1946
Qy 1801 AGACCCAGATCACAAAAGCCAAACCATGCCAGCCCTCCAGCACCCCAAGCCCAAGAC 1860
Db 1947 AGACCCAGATCACAAAAGCCAAACCATGCCAGCCCTCCAGCACCCCAAGCCCAAGAC 2006
Qy 1861 CATGTTCTGAATTCTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
Db 2007 CATGTTCTGAATTCTGACGACACTGTGAGCCCTTGTACTTTAACTCATGGAAG 2066
Qy 1921 ATAACCTCTCAGCTTTGAATAATGTTTCTCTGTGAA 1962
Db 2067 ATAACCTCTCAGCTTTGAATAATGTTTCTCTGTGAA 2108

RESULT 7
AX775943
LOCUS AX775943 2304 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 213 from Patent WO03048202.
ACCESSION AX775943
VERSION AX775943.1 GI:32693661
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 213 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1. 2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362. 1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB11671.1"
/translation="MDVVEVAGSMWAQEREDIMKYEKHRAGLPEDKPKFRSYNN
NVDHLGIVHEBELPPLTAREAKQIREISRSKMVDMLDWEKYKSSRKLIDRAYKGM
PMNIRGPMMSVLLNTEBMKLNPRRYQIMKEKGRSSEHIQRIDRVSGTLRKHIFPR
DRYGTQKRELHILILAYERVNPVEGCRDLSHIALFLYLPEEDAFWALVOLLASER
HSLQGFHSPNGSTVQLODQOEHVAVTSQPKTMGHQDKDLGQCSPLGCLIRILIDG
ISLGLTRLMVDVYLVEGEOLMPITRIAFAKVOQKRLTTSRCPWAFNRFVDTMAR
DEDTVLKHLRASMKLTKRQGDLPPEAKPEQSSASRPVPSRGGKTLCKGDRQAPPG
PFGQLPRHGLVLPVPLVGLSGKPTLWALRVCPARPWLREDLRVPGDSCSTPCPA
SORTWT"
ORIGIN
Query Match 96.8%; Score 1902; DB 6; Length 2304;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1937; Conservative 0; Mismatches 5; Indels 22; Gaps 1;
Qy 1 ATGGAAGTGTAGAGTTCGGGCAAGTTGTGGGCAACAAGAGCGAGAGACATCATTTATG 60
Db 362 ATGGAAGTGTAGAGTTCGGGCAAGTTGTGGGCAACAAGAGCGAGAGACATCATTTATG 421

Qy	61	AAATACGAAAAGGACA	CGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCGA	120		
Db	422	AAATACGAAAAGGACA	CGAGCTGGGCTGCCAGAGCAAGGGGCTTAAGCTTTTCGA	481		
Qy	121	AGCTACAACAACACGTCGATCATTTGGGATTTGTACATGAGACGGAGCTGCTCTG	180			
Db	482	AGCTACAACAACA	CGTCGATCATTTGGGATTTGTACATGAGACGGAGCTGCTCTG	541		
Qy	181	ACTCGCGGGAGGCGAAGCAA	TTCCGCGGAGATCAGCCGAAAGAACAAGTGGGTGAT	240		
Db	542	ACTCGCGGGAGGCGAAGCAA	TTCCGCGGAGATCAGCCGAAAGAACAAGTGGGTGAT	601		
Qy	241	ATGCTGGGAGACTGGAGAAATACA	AAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG	300		
Db	602	ATGCTGGGAGACTGGAGAAATACA	AAAGCAGCAGAAAGCTCATAGATCGAGCGTACAAG	661		
Qy	301	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCA	GTCCTCTGAACATTGAGGAAATG	360		
Db	662	GGAATGCCCATGAACATCCGGGGCCCGATGTGTCA	GTCCTCTGAACATTGAGGAAATG	721		
Qy	361	AAGTGAAAAA	CCCCGGAAAGATACAGATCATGAAGAGAAAGGCAAGAGTCATCTGAG	420		
Db	722	AAGTGAAAAA	CCCCGGAAAGATACAGATCATGAAGAGAAAGGCAAGAGTCATCTGAG	781		
Qy	421	CACATCCAGCGCATCGACCGGGA	CGTAAACGGGACATTAAAGAAAGCATATATTTCTTCAGG	480		
Db	782	CACATCCAGCGCATCGACCGGGA	CGTAAACGGGACATTAAAGAAAGCATATATTTCTTCAGG	841		
Qy	481	GATCGATACGGAACCAAGCA	CGGGAACTACTCCACATCTCTTGCCATATGAGGAGTAT	540		
Db	842	GATCGATACGGAACCAAGCA	CGGGAACTACTCCACATCTCTTGCCATATGAGGAGTAT	901		
Qy	541	AAACCCGAGGTGGGCTACTGCA	GGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT	600		
Db	902	AAACCCGAGGTGGGCTACTGCA	GGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT	961		
Qy	601	CTTCCTGAGGAGGATGATTTCTGGGCA	CTGTGACGTGCTGGCCAAGTAGAGGCACTCC	660		
Db	962	CTTCCTGAGGAGGATGATTTCTGGGCA	CTGTGACGTGCTGGCCAAGTAGAGGCACTCC	1021		
Qy	661	CTGCAAGGATTTCA	CAGCCCAATGGCGGGAACCGTCCAGGGGCTCCAAGACCAACAGAG	720		
Db	1022	CTGCAAGGATTTCA	CAGCCCAATGGCGGGAACCGTCCAGGGGCTCCAAGACCAACAGAG	1081		
Qy	721	CATGTGTAAGCCACGTCA	CAACCCCAAGACCATGSGGCAATCAAGCAAGAAGATCTATGT	780		
Db	1082	CATGTGTAAGCCACGTCA	CAACCCCAAGACCATGSGGCAATCAAGCAAGAAGATCTATGT	1141		
Qy	781	GGGCAGTGTTCCCGCTTA	GGCTGAGCTCATCCGATATTTGAACGGGATCTCTCTCGGG	840		
Db	1142	GGGCAGTGTTCCCGCTTA	GGCTGAGCTCATCCGATATTTGAACGGGATCTCTCTCGGG	1201		
Qy	841	CTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAA	GGCGAACAGGCGTTGATGCCGATA	900		
Db	1202	CTCACCCCTGCGCTGTGGGACGTGTATCTGTAGAA	GGCGAACAGGCGTTGATGCCGATA	1261		
Qy	901	ACAAGAATCGCCTTTAAGTT	CAGCAGAAGCGCTCAGAAAGCGTCCAGGTGTGCCCG	960		
Db	1262	ACAAGAATCGCCTTTAAGTT	CAGCAGAAGCGCTCAGAAAGCGTCCAGGTGTGCCCG	1321		
Qy	961	TGGGCAAGTTTGTGCA	ACCGGTTGTTGATACCTGGGCGCAGGGA	TGAGGACA	CTGTGCTC	1020
Db	1322	TGGGCAAGTTTGTGCA	ACCGGTTGTTGATACCTGGGCGCAGGGA	TGAGGACA	CTGTGCTC	1381
Qy	1021	AAGCATCTTAAAGGCTCTAT	GAAGAACTAAACAAGAAAGCAGGGGGA	ACCTGCCACCCCA		1080
Db	1382	AAGCATCTTAAAGGCTCTAT	GAAGAACTAAACAAGAAAGCAGGGGGA	ACCTGCCACCCCA		1441
Qy	1081	GCCAAACCCGAGCAAGGGT	CGTCGGCATCCAGGCGCTGTGCCGCTTCA	CGTGGCGGAAG		1140
Db	1442	GCCAAACCCGAGCAAGGGT	CGTCGGCATCCAGGCGCTGTGCCGCTTCA	CGTGGCGGAAG		1501
Qy	1141	ACCCTCTGCAAGGGGGA	CAGGCAAGGCCCTCCAGGCCCAACG	CCCCGGTTCCCGCGGCC		1200

Db	1502	ACCCCTGCAAGGGGGACAGGCAAGCCCTTCA-----GGCCC	1539
QY	1201	ATTGGTCAAGCTCCCCGCAAGGCACTCGTTCTTCCACACCCTGCTGTGGGCT	1260
Db	1540	ATTGGTCAAGCTCCCCGCAAGGCACTCGTTCTTCCACACCCTGCTGTGGGCT	1599
QY	1261	GTCGGGAAGACACCTACCCTGTGGGCACTCAGGGGTGTGCCAGCCCGCCCTGGCTCAG	1320
Db	1600	GTCCGGGAAGACACCTACCCTGTGGGCACTCAGGGGTGTGCCAGCCCGCCCTGGCTCAG	1659
QY	1321	GGAGGACCTCAGGGTTCCTGGAGATTCTGCACTGGAACCTCCATGCCCGCTCCCAACG	1380
Db	1660	GGAGGACCTCAGGGTTCCTGGAGATTCTGCACTGGAACCTCCATGCCCGCTCCCAACG	1719
QY	1381	GACCTGGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTACAGACAGAGCTGTGGGTC	1440
Db	1720	GACCTGGACGTAGAGGGCCCTGTGTTCCGCCATTATGATTTACAGACAGAGCTGTGGGTC	1779
QY	1441	CGTGCCATATCCCAAGAGGACCAAGCTGAGCCCTGCTGGCAGGCTGAACACCCTGCGGAG	1500
Db	1780	CGTGCCATATCCCAAGAGGACCAAGCTGAGCCCTGCTGGCAGGCTGAACACCCTGCGGAG	1839
QY	1501	CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGGAACCCCTTCAGA	1560
Db	1840	CGGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACACAGGGAACCCCTTCAGA	1899
QY	1561	GCTAGGGAAGAACAGCCCGTGTGTCTCCCACTCAGGGCCCTTGCCCTCGCGGCTCCACTTG	1620
Db	1900	GCTAGGGAAGAACAGCCCGTGTGTCTCCCACTCAGGGCCCTTGCCCTCGCGGCTCCACTTG	1959
QY	1621	GAAAGTCTCAGTTCCTCCCTCAGGCTTCTAAGAACATCTGGGCCAGGGCTCATGGCTGGAT	1680
Db	1960	GAAAGTCTCAGTTCCTCCCTCAGGCTTCTAAGAACATCTGGGCCAGGGCTCATGGCTGGAT	2019
QY	1681	AATTTCCCTAGGCTTAAACAACCCAGCAAGCTTGCGCTCTCGTTTATTTTGGTTAAA	1740
Db	2020	AATTTCCCTAGGCTTAAACAACCCAGCAAGCTTGCGCTCTCGTTTATTTTGGTTAAA	2079
QY	1741	CTTATGAAATGTATTAAGAAAGATGACCTCGAGAGATTACAGAGTGAACACACC	1800
Db	2080	CTTATGAAATGTATTAAGAAAGATGACCTCGAGAGATTACAGAGTGAACACACACC	2139
QY	1801	AGACCCAGATCACAAAGCCAACCATGCCCCAGGCCCTCCAGCACCCCCAGGCCCAAGC	1860
Db	2140	AGACCCAGATCACAAAGCCAACCATGCCCCAGGCCCTCCAGCACCCCCAGGCCCAAGC	2199
QY	1861	CATCGTCTGAATTCGACGACACACCGTGAGCTTGCTTGTACTTTAACTCATGGAAG	1920
Db	2200	CATCGTCTGAATTCGACGACACACCGTGAGCTTGCTTGTACTTTAACTCATGGAAG	2259
QY	1921	ATAACTACCTTCACGTTTGAATAAATGTTCTGTGAAATG	1964
Db	2260	ATAACCACCTTCATGTTTGAATAAATGTTCTGTGAAATG	2303

RESULT 8			
AX775941			
LOCUS	AX775941	2304 bp	mRNA
DEFINITION	Sequence 211 from Patent WO03048202.	linear	PAT 14-JUL-2003
ACCESSION	AX775941		
VERSION	AX775941.1	GI:32693659	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1		
AUTHORS	Matsuda, A. and Muramatsu, S.		
TITLE	NF-kB activating gene		
JOURNAL	Patent: WO 03048202-A 211 12-JUN-2003;		
	Asahi Kasei Kabushiki Kaisha (JP)		

FEATURES
source
location/Qualifiers
1..2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11670.1"
/db_xref="GI:32693660"
translation="MDVVEVAGSWMAQEREDIMKYEKGRAGLPEDKPKPFRSYNNVDHLGIVHETELPPLTAREAKQIRREISRSKWVDMLGDWEKYSRXLIDRAYKM
PMNIRGPMMSVLNLTENMKLNPGRYQIMKEKRSSEHIIQIRIDRVSGTLRKHIFPR
DRYGTQRELIHLILAYEYNEVGYCRDLSHIALFLYLPEBDAFWALVQLASER
HSLQGFHSPNGTVOGLDQDQEHVATSQPTMGHQDKDLGGCCSPLGCLIRILIDG
ISLGLTLRLMDVYLVEGSOALMPITRIAFKVOQKRLLTKTSRCGPMARFCNRFVDTWAR
DEDTVLKHLRASMKLTRKQGLQPPAKPEQSSASRPVPASRGKTLCKGDRQAPG
PEGQLRHHGLVLPHPVLVGLSGKPTLWALRVCPARPWLREDLRVPGDSCSGTPCPA
SORTWT"

ORIGIN

Query Match 96.8%; Score 1900.4; DB 6; Length 2304;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1936; Conservative 0; Mismatches 6; Indels 22; Gaps 1;

QY 1 ATGACGTGCTAGAGGTGCGCGGCACTGTGGGCACAGAAGCGAGACATCATTTATG 60
DB 362 ATGACGTGCTAGAGGTGCGCGGCACTGTGGGCACAGAAGCGAGACATCATTTATG 421
QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTAAGCCTTTTCCA 120
DB 422 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTAAGCCTTTTCCA 481
QY 121 AGCTACACAACAACGTGATCATTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG 180
DB 482 AGCTACACAACAACGTGATCATTTGGGGATTGTACATGAGACGGAGCTGCCTCTCTG 541
QY 181 ACTGCGCGGGAAGCGAAGCAATTCGGCGGAGATCAGCCGAAAGACCAAGTGGGTGAT 240
DB 542 ACTGCGCGGGAAGCGAAGCAATTCGGCGGAGATCAGCCGAAAGACCAAGTGGGTGAT 601
QY 241 ATGCTGGGAGACTGGGAAATACAAAAGCAGCAGAAAGCTCATAGTCGAGCTACAAG 300
DB 602 ATGCTGGGAGACTGGGAAATACAAAAGCAGCAGAAAGCTCATAGTCGAGCTACAAG 661
QY 301 GGAATGCCATGAACATCCGGGCGCGATGTGTCACTCTCTGAACATTTAGGAAATG 360
DB 662 GGAATGCCATGAACATCCGGGCGCGATGTGTCACTCTCTGAACATTTAGGAAATG 721
QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCACTGAG 420
DB 722 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGAAAGGCAAGGTCACTGAG 781
QY 421 CACATCCAGCGCATCGACCCGGACGTAAGCGGACATTAAGGACATATATTCTTCAGG 480
DB 782 CACATCCAGCGCATCGACCCGGACGTAAGCGGACATTAAGGACATATATTCTTCAGG 841
QY 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 540
DB 842 GATCGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATGAGAGTAT 901
QY 541 AACCCGAGGTGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
DB 902 AACCCGAGGTGGCTACTGACGGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 961
QY 601 CTTCTGAGGAGATGATTTCTGGGCACTGTGCACTGCTGGCCAGTGAAGGCACTCC 660
DB 962 CTTCTGAGGAGATGATTTCTGGGCACTGTGCACTGCTGGCCAGTGAAGGCACTCC 1021
QY 661 CTGACGGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
DB 1022 CTGACGGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 1081

QY 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAAGACAAGAAAGATCTATGT 780
DB 1082 CATGTGTAGCCACGTCACAAACCAAGACCATGGGGCATCAAGACAAGAAAGATCTATGT 1141
QY 781 GGGCAGTGTTCGCCGTTAGGCTGCCTCATCCGGAATAATTGACGGGATCTCTCCGG 840
DB 1142 GGGCAGTGTTCGCCGTTAGGCTGCCTCATCCGGAATAATTGACGGGATCTCTCCGG 1201
QY 841 CTCACCTGCGCCTGTGGGACGTGTATCTGTGAGAGCGCAACAGGCGTTGATGCCGATA 900
DB 1202 CTCACCTGCGCCTGTGGGACGTGTATCTGTGAGAGCGCAACAGGCGTTGATGCCGATA 1261
QY 901 ACAAGAAATCGCCTTTAAGGTTCAGCAGAGAGCGCCTCAAGAGACGTCCAGGTGGCCCG 960
DB 1262 ACAAGAAATCGCCTTTAAGGTTCAGCAGAGAGCGCCTCAAGAGACGTCCAGGTGGCCCG 1321
QY 961 TGGGCACTTTTGCAAACCGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1020
DB 1322 TGGGCACTTTTGCAAACCGTTCTGTGATACCTGGGCCAGGATGAGACACTGTGCTC 1381
QY 1021 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCACCCCA 1080
DB 1382 AAGCATCTTAGGGCTCTATGAAGAACTAACAAAGAAAGCAGGGGACCTGCCACCCCA 1441
QY 1081 GCCAAACCCGACGAAGGTCCTCGGCATCCAGGCTGTGCCGCTTCACTGCGCGGAAG 1140
DB 1442 GCCAAACCCGACGAAGGTCCTCGGCATCCAGGCTGTGCCGCTTCACTGCGCGGAAG 1501
QY 1141 ACCCTCTGCAAGGGGACAGGAGGCCCTCCAGGCCCAACAGCCCGGTTCCGCGGCC 1200
DB 1502 ACCCTCTGCAAGGGGACAGGAGGCCCTCCCA-----GGCCC 1539
QY 1201 ATTGTGACCTTCCCGCCACCGGCACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1260
DB 1540 ATTGTGACCTTCCCGCCACCGGCACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1599
QY 1261 GTCCGGAAAGACACCTACCTGTGGGCACTCAGGCTGTCCAGCCCGGCTGGCTCAG 1320
DB 1600 GTCCGGAAAGACACCTACCTGTGGGCACTCAGGCTGTCCAGCCCGGCTGGCTCAG 1659
QY 1321 GGAAGACCTCAGGTTCTGTGAGATTCTCTGACGTGAATCTCATGCCCGCTCCCAAG 1380
DB 1660 GGAAGACCTCAGGTTCTGTGAGATTCTCTGACGTGAATCTCATGCCCGCTCCCAAG 1719
QY 1381 GACCTGGAAGTGAAGGGCCCTTGTGTTCCGCCATTATGATTTCAAGCAGAGCTGTGGTC 1440
DB 1720 GACCTGGAAGTGAAGGGCCCTTGTGTTCCGCCATTATGATTTCAAGCAGAGCTGTGGTC 1779
QY 1441 CGTGCCATATCCAGGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1500
DB 1780 CGTGCCATATCCAGGAGACCAAGCTGGCCCCCTGTGGCAGGCTGAACACCTGCGGAG 1839
QY 1501 CGGTGAGATCGGCTTTGCTGCAACCCAGCAGTATCCGACAGGCGACCCCTTCAGA 1560
DB 1840 CGGTGAGATCGGCTTTGCTGCAACCCAGCAGTATCCGACAGGCGACCCCTTCAGA 1899
QY 1561 GCTAGGAGCGAACAGCCGCTGTCTCCCACTCAGGCTTGCTCTGCGGCTCCACTTG 1620
DB 1900 GCTAGGAGCGAACAGCCGCTGTCTCCCACTCAGGCTTGCTCTGCGGCTCCACTTG 1959
QY 1621 GAAAGTTCTCAGTTCCTTCAGGCTTCTAGAAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680
DB 1960 GAAAGTTCTCAGTTCCTTCAGGCTTCTAGAAAGCATCTGGGCCAGGGCTCATGGCTGAT 2019
QY 1681 AATTTCCTAGGCTTAAACAACCAAGCAAGCTTCGCTCTGTTTATTGTTGTTAAA 1740
DB 2020 AATTTCCTAGGCTTAAACAACCAAGCAAGCTTCGCTCTGTTTATTGTTGTTAAA 2079
QY 1741 CTTATGAAAAATGTATTAAGAAAGAGTGACGCTGAGAGAGATTCAAGATGGAACACACC 1800
DB 2080 CTTATGAAAAATGTATTAAGAAAGAGTGACGCTGAGAGAGATTCAAGATGGAACACACC 2139
QY 1801 AGACCCAGATCAAAAGCCAACCATGCCCCGCTCCCAAGACCCCAAGCCCAAGAC 1860

Db 2140 AGACCCAGATCACAAAGCCAAATGCCCCAGCCCCCAGACACCCCAAGCCCAAGCAG 2199
Qy 1861 CATCGTTCTGAATTTCTGACGACACCGTGAGCCCTGCTTTGTACTTTAACTCATGGAAG 1920
Db 2200 CATCGTTCTGAATTTCTGACGACACCGTGAGCCCTGCTTTGTACTTTCAACTCATGGAAG 2259
Qy 1921 ATAACCTCTTCACGTTTGAATAATGTTTCTGTTGAATG 1964
Db 2260 ATAACCACTTCATGTTTGAATAATGTTTCTGTTGAATG 2303

RESULT 9
AR578516
LOCUS AR578516 2084 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 692 from patent US 6783969.
ACCESSION AR578516
VERSION AR578516.1 GI:56581312
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2084)
AUTHORS Tang,Y.T., Goodrich,R.W., Asundi,V. and Dirmanac,R.T.
TITLE Cathepsin V-like polypeptides
JOURNAL Patent: US 6783969-A 692 31-AUG-2004;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
source 1..2084
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 93.5%; Score 1836.4; DB 6; Length 2084;
Best Local Similarity 96.9%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 6; Indels 54; Gaps 1;

Qy 1 ATGACGTGTAGAGGTGCGCGGCAAGTTGGTGACAAAGAGAGAGACATCATATG 60
Db 163 ATGACGTGTAGAGGTGCGCGGCAAGTTGGTGACAAAGAGAGAGACATCATATG 222
Qy 61 AAATACGAAAGGGACACCGAGCTGGCTGCCAGAGACAAGGGCCCTAAGCCTTTTGA 120
Db 223 AAATACGAAAGGGACACCGAGCTGGCTGCCAGAGACAAGGGCCCTAAGCCTTTTGA 282
Qy 121 AGCTACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 180
Db 283 AGCTACAACAACGTCGATCATTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 342
Qy 181 ACTGCGCGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAAGACAAAGTGGTGAT 240
Db 343 ACTGCGCGAGGCGAAGCAAAATTCGCGGAGATCAGCCGAAAGACAAAGTGGTGAT 402
Qy 241 ATGCTGGGAGACTGGAGAAATACAAAGCAGAGAAAGCTCATAGTCGAGCGTACAAG 300
Db 403 ATGCTGGGAGACTGGAGAAATACAAAGCAGAGAAAGCTCATAGTCGAGCGTACAAG 462
Qy 301 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACATTGAGAAATG 360
Db 463 GGAATGCCATGAACATCCGGGGCCGATGTGTCACTCTCTGAACACTGAGAAATG 522
Qy 361 AAGTTGAAAAAAGCCCGGAAGATACGATCATGAAGAGAAAGGCAAGGTCTGTAG 420
Db 523 AAGTTGA----- 529
Qy 421 CACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGAAAGCATATATTCTTCAG 480
Db 530 -ACATCCAGCGCATCGACCGGACGTAAAGCGGACATTAAGAAAGCATATATTCTTCAG 588
Qy 481 GATGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTTGACATATGAGAGTAT 540
Db 589 GATGATACGGAACCAAGCAGCGGAACTACTCCACATCTCTTGACATATGAGAGTAT 648

Qy 541 AACCCGAGGTGGGCTACTGACAGGAGCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 649 AACCCGAGGTGGGCTACTGACAGGAGCTGAGCCACATCGCCGCTTGTCTCTCTAT 708
Qy 601 CTTCTGAGGAGGATGCATTTCTGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTCC 660
Db 709 CTTCTGAGGAGGATGCATTTCTGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTCC 768
Qy 661 CTGAGGGATTTCACAGCCCAATGGCGGAGCCGTCAGGGGCTCCAAGACCAACAGAG 720
Db 769 CTGAGGGATTTCACAGCCCAATGGCGGAGCCGTCAGGGGCTCCAAGACCAACAGAG 828
Qy 721 CATGTGTAGCCACGTCAACCCCAAGACCATGGGCGATCAGACAAGAAAGATCTATGT 780
Db 829 CATGTGTAGCCACGTCAACCCCAAGACCATGGGCGATCAGACAAGAAAGATCTATGT 888
Qy 781 GGGCAGTGTCCCGTTAGGCTGCTCATCCGGAATGATTGACGGGATCTCTCGGG 840
Db 889 GGGCAGTGTTCGCCGTTAGGCTGCTCATCCGGAATGATTGACGGGATCTCTCGGG 948
Qy 841 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAAAGCGGAACAGGCGTTGATGCCGATA 900
Db 949 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAAAGCGGAACAGGCGTTGATGCCGATA 1008
Qy 901 ACAAGATCGCCTTTAAGTTCAAGCAAGAGCGCTCAAGAAAGCTCCAGGTGCGCCG 960
Db 1009 ACAAGATCGCCTTTAAGTTCAAGCAAGAGCGCTCAAGAAAGCTCCAGGTGCGCCG 1068
Qy 961 TGGGCACTTTTTCACACCGGTTCGTTGATACCTGCGCCAGGAGTAGAGACATGTGCTC 1020
Db 1069 TGGGCACTTTTTCACACCGGTTCGTTGATACCTGCGCCAGGAGTAGAGACATGTGCTC 1128
Qy 1021 AAGCATTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGAGCTGCCACCCCA 1080
Db 1129 AAGCATTTAGGGCTCTATGAAGAACTAACAGAAAGCAGGGGAGCTGCCACCCCA 1188
Qy 1081 GCCAAACCCGAGCAAGGGTGTGGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG 1140
Db 1189 GCCAAACCCGAGCAAGGGTGTGGCATCCAGGCTGTGCCGCTTCAAGTGGCGGAG 1248
Qy 1141 ACCCTGTCAAGGGGAGACAGGCAAGCCCTCCAGGCCCAACAGCCCGGTTCCCGGCC 1200
Db 1249 ACCCTGTCAAGGGGAGACAGGCAAGCCCTCCAGGCCCAACAGCCCGGTTCCCGGCC 1308
Qy 1201 ATTGTGAGCTTCCCGCCACGGGCACTCTGTTCTTCACACCTGTCTGTGGGCT 1260
Db 1309 ATTGTGAGCTTCCCGCCACGGGCACTCTGTTCTTCACACCTGTCTGTGGGCT 1368
Qy 1261 GTCCGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1320
Db 1369 GTCCGGAAGACACTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGCTCAG 1428
Qy 1321 GGAAGACTCAGGGTTCGTGAGATTCTGCAGTGAACTCCATGCCCGCTCCCAAG 1380
Db 1429 GGAAGACTCAGGGTTCGTGAGATTCTGCAGTGAACTCCATGCCCGCTCCCAAG 1488
Qy 1381 GACCTGSAAGTGAAGGGCCCTTGGTTCCGCCATTAATTTCAAGACAGAGCTGTGGTC 1440
Db 1489 GACCTGSAAGTGAAGGGCCCTTGGTTCCGCCATTAATTTCAAGACAGAGCTGTGGTC 1548
Qy 1441 CGTGCCATATCCAGGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1500
Db 1549 CGTGCCATATCCAGGAGACAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCGGAG 1608
Qy 1501 CGGGTAGATCGGCTTTCGTGACCCAGACCTGAATTCGACAGGCAAGCCCTTCAGA 1560
Db 1609 CGGGTAGATCGGCTTTCGTGACCCAGACCTGAATTCGACAGGCAAGCCCTTCAGA 1668
Qy 1561 GCTAGGACGAACAGCCGTGTCTCCACCTCAGGGCTTGCTCTGCGGCTTCACTTG 1620
Db 1669 GCTAGGACGAACAGCCGTGTCTCCACCTCAGGGCTTGCTCTGCGGCTTCACTTG 1728
Qy 1621 GAAAGTTCTAGTTCCTCCAGGCTTCTAGAAAGCATCTGGGCCAGGGCTCATGGCTGAT 1680

Db 1729 GAAAGTCTCAGTTCCTCCAGGCTTTAGAGCATCTGGGCCAGGGCTCATGGCTGCAT 1788
Qy 1681 AATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGGCTCTGTTTATTTTGGTTAAA 1740
Db 1789 AATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGCATCTCGTTTATTTTGGTTAAA 1848
Qy 1741 CTTATGAAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACAACC 1800
Db 1849 CTTATGAAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACAACC 1908
Qy 1801 AGACCCAGATCACAAAAGCCAAAGCCAGCCCTCCAGCAACCCCAAGCCCAAGAC 1860
Db 1909 AGACCCAGATCACAAAAGCCAAAGCCAGCCCTCCAGCAACCCCAAGCCCAAGAC 1968
Qy 1861 CATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGAAAG 1920
Db 1969 CATCGTTCTGAATTTCTGACGACACCGTGAGCCTGCTTGTACTTTCAAACTCATGAAAG 2028
Qy 1921 ATAACCTACCTTCACGTTTGAAATTAATGTTTCTGTTGAAATG 1964
Db 2029 ATAACCACTTCATGTTTGAATTAATGTTTCTGTTGAAATG 2072

RESULT 10
AR578515 2072 bp DNA linear PAT 14-DEC-2004
LOCUS AR578515 Sequence 691 from patent US 6783969.
DEFINITION AR578515
ACCESSION AR578515
VERSION AR578515.1 GI:56581311
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Tang,Y.T., Goodrich,R.W., Asundi,V. and Dirmnac,R.T.
TITLE Cathepsin V-like polypeptides
JOURNAL Patent: US 6783969-A 691 31-AUG-2004;
Nuvelo, Inc.; Sunnyvale, CA
FEATURES
location/Qualifiers
source 1..2072
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 92.3%; Score 1812.4; DB 6; Length 2072;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1892; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
Qy 1 ATGACGTGTGTAAGGTGCGGGGCACTTGTGGGCAACAAGCGAGACATCATATATG 60
Db 163 ATGACGTGTGTAAGGTGCGGGGCACTTGTGGGCAACAAGCGAGACATCATATATG 222
Qy 61 AAATAGCAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 120
Db 223 AAATAGCAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTGA 282
Qy 121 AGCTACAACAACAACGTGCATCTTGGGATTGTACATGAGACGGAGCTGCTCTCTG 180
Db 283 AGCTACAACAACAACGTGCATCTTGGGATTGTACATGAGACGGAGCTGCTCTCTG 342
Qy 181 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 240
Db 343 ACTGCGGGGAGGCGAAGCAAAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 402
Qy 241 ATGCTGGGAGACTGGAGAAATACAAGACAGCAAGAACTCATAGATCGAGCTACAAG 300
Db 403 ATGCTGGGAGACTGGAGAAATACAAGACAGCAAGAACTCATAGATCGAGCTACAAG 462
Qy 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTCTCACTCTCTGAACATTTAGGAAATG 360
Db 463 GGAATGCCCATGAACATCCGGGGCCGATGTGTCTCACTCTCTGAACATTTAGGAAATG 522

Qy 361 AAGTTGAAAAAAGCCGGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 420
Db 523 AAGTTGAAAAAAGCCGGAAGATACAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 582
Qy 421 CACATCCAGCGCATCGACCCGGGACGTAAGCGGGAATTAAGGAAGCATATATCTTCAAG 480
Db 583 CACATCCAGCGCATCGACCCGGGACGTAAGCGGGAATTAAGGAAGCATATATCTTCAAG 642
Qy 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTGGCATATGAGAGTAT 540
Db 643 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCACATCTCTGGCATATGAGAGTAT 702
Qy 541 AACCAGAGGTGGGCTACTGAGAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 600
Db 703 AACCAGAGGTGGGCTACTGAGAGGACCTGAGCCACATCGCCGCTTGTCTCTCTAT 762
Qy 601 CTTCTGAGGAGATGATTCATCTGGGCACCTGTGCAGCTGCGCCAGTGAAGGCACTCC 660
Db 763 CTTCTGAGGAGATGATTCATCTGGGCACCTGTGTGAGCTGTGGCCAGTGAAGGCACTCC 822
Qy 661 CTGAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
Db 823 CTGAGGATTTCAAGCCCAATGGCGGACCGTCCAGGGGCTCCAAGACCAAGAGAG 882
Qy 721 CATGTGTAGCCACGTCACAACCCAAAGACCATGCGGCAATCAAGACAAAGAAATCTATGT 780
Db 883 CATGTGTAGCCACGTCACAACCCAAAGACCATGCGGCAATCAAGACCAAGAGAG 923
Qy 781 GGGCAGTGTTCCTGAGGCTGCTCATCCGATATTTGATGACGGGATCTCTCGGG 840
Db 924 -----GATCTCTCTCGGG 936
Qy 841 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGGCGTTGATGCCGATA 900
Db 937 CTCACCTGCGCCTGTGGAGCTGTATCTGTAGAAAGCGAAACAGGCGTTGATGCCGATA 996
Qy 901 ACAAGATCGCCTTTAAGTTCAAGAGAGCGCTCAAGAAAGCTCCAGGTGCGCCG 960
Db 997 ACAAGATCGCCTTTAAGTTCAAGAGAGCGCTCAAGAAAGCTCCAGGTGCGCCG 1056
Qy 961 TGGGCACTTTTTCACAACCGGTTCTGTATACCTGCGCAAGGATGAGGACACTGTGCTC 1020
Db 1057 TGGGCACTTTTTCACAACCGGTTCTGTATACCTGCGCAAGGATGAGGACACTGTGCTC 1116
Qy 1021 AAGCATCTTAGGGCTCTATGAAGAACTTAACAAGAAAGCAAGGGGACCTGCCACCCCA 1080
Db 1117 AAGCATCTTAGGGCTCTATGAAGAACTTAACAAGAAAGCAAGGGGACCTGCCACCCCA 1176
Qy 1081 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCGGCTTCACTGCGCGGAAG 1140
Db 1177 GCCAAACCCGAGCAAGGGTGTGCGCATCCAGGCTGTGCGGCTTCACTGCGCGGAAG 1236
Qy 1141 ACCCTCTGCAAGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
Db 1237 ACCCTCTGCAAGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1296
Qy 1201 ATTTGTCAGCTTCCCGGCAAGGACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1260
Db 1297 ATTTGTCAGCTTCCCGGCAAGGACCTCGTTCTTCCACACCTGTCTGTGGGGCT 1356
Qy 1261 GTCCGGGAAGACACCTTACCTGTGGGACACTCAGGAGTGTCCCAAGCCCGGCTGGCTCAG 1320
Db 1357 GTCCGGGAAGACACCTTACCTGTGGGACACTCAGGAGTGTCCCAAGCCCGGCTGGCTCAG 1416
Qy 1321 GAGGACCTCAGGGTTCTGAGAGATTCTGCACTGGAATCTCATGCCCCGCTCCCAAG 1380
Db 1417 GAGGACCTCAGGGTTCTGAGAGATTCTGCACTGGAATCTCATGCCCCGCTCCCAAG 1476
Qy 1381 GACCTGACGTAGAGGGCTTGTGCTCGCCATATGATTTAGACAGAGCTGTGGGTC 1440
Db 1477 GACCTGACGTAGAGGGCTTGTGCTCGCCATATGATTTAGACAGAGCTGTGGGTC 1536
Qy 1441 CGTGCCATATCCAGAGAGACAGCTGAGCCCTGCTGGCAGGCTGAACACCTGCGAG 1500

Db 1537 CGTGCATATCCAGAGAGACCAGCTGGCCCCCTGCTGGCAGGCTGAACACCCTGGCAG 1596
Qy 1501 CGGGTGAATCGGCTTTGCGTGCACCCGACACTGATTTCCGACAGGGGCACCCCCTTACA 1560
Db 1597 CGGGTGAATCGGCTTTGCGTGCACCCGACACTGATTTCCGACAGGGGCACCCCCTTACA 1656
Qy 1561 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCCCTCTGCGGCTCCACTTG 1620
Db 1657 GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCCCTCTGCGGCTCCACTTG 1716
Qy 1621 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT 1680
Db 1717 GAAAGTTCAGTTCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGGCTGGAT 1776
Qy 1681 AATTCCCTAGGCTTAAACAACCCAGCAAGCTTCGCGTCTGCTTTTATTGTTAA 1740
Db 1777 AATTCCCTAGGCTTAAACAACCCAGCAAGCTTCGCACTCCTGTTTATTGTTAA 1836
Qy 1741 CTTATGAAAATGTATTAGAAGAGTGCAGCTCGAGAGATTCAGAGATGAACACACC 1800
Db 1837 CTTATGAAAATGTATTAGAAGAGTGCAGCTCGAGAGATTCAGAGATGAACACACC 1896
Qy 1801 AGACCCAGATCACAAAGCCAAACCATGCCCCCTCCAGCACCCCCAAGCCACGAC 1860
Db 1897 AGACCCAGATCACAAAGCCAAACCATGCCCCCTCCAGCACCCCCAAGCCACGAC 1956
Qy 1861 CATCGTTCGAATTTCTGACGACACCGGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
Db 1957 CATCGTTCGAATTTCTGACGACACCGGTGAGCCTGCTTGTACTTTCAAACTCATGGAAG 2016
Qy 1921 ATAACTACCTTCACGTTTGAATTAATGTTTCTGTTGAAATG 1964
Db 2017 ATAACTACCTTCATGTTTGAATTAATGTTTCTGTTGAAATG 2060

RESULT 11
AX75951 2647 bp mRNA linear PAT 14-JUL-2003
LOCUS AX75951 Sequence 221 from Patent WO03048202.
ACCESSION AX75951
VERSION AX75951.1 GI:32693669
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
AUTHORS Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 221 12-JUN-2003;
Asahi Kabushiki Kaisha (JP)
FEATURES
Source Location/Qualifiers
1. 2647
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11675.1"
/db_xref="GI:32693670"
/translation="MDVVEVAGSWWAQEREDIMKYEKHRAGLPEDKBEPFRSYNN
NVDHGI VHTBLPLTAREAKQIR EISRSKMYDMLGDMWKYSRKLIDRAYKM
PMNIRGPMVSLNTEEMKLNKPGRYQIMKEKGRSEH IQRIDRVSGTLRKHIIFR
DRYGTQRRLHLILA YEBNPEVGYCRDLSHIALFLYLPEEDA FMAVLQLLASER
HSLQGFHSPNGTVQGLDQOQEHVATSQPKTMGQISLGLTLRLWDVYLVEGOALM
PTRIAFKVQKRLTKTSRSGPMARFCNRFVDTWARDEDTVLKHLSAKKLTTRKQD
LPPAKPEQSSASRPVPASRGKTLCKGDRAPGPPARPRPIWSASPPRAPRSST
PCPGAVREDTYPVGTQGVPSPALAQGGPQGSWRFLONSMPRLPTDL DVEGPMFRHY
DFROS CMVRAISOEDQLAPCWAQAEHPAERVRSAFAPASTDS DQTFPRARDEQPCAPT
SGPCLCGHLBESSQFP PGP"

CDS
755..2338
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11675.1"
/db_xref="GI:32693670"
/translation="MDVVEVAGSWWAQEREDIMKYEKHRAGLPEDKBEPFRSYNN
NVDHGI VHTBLPLTAREAKQIR EISRSKMYDMLGDMWKYSRKLIDRAYKM
PMNIRGPMVSLNTEEMKLNKPGRYQIMKEKGRSEH IQRIDRVSGTLRKHIIFR
DRYGTQRRLHLILA YEBNPEVGYCRDLSHIALFLYLPEEDA FMAVLQLLASER
HSLQGFHSPNGTVQGLDQOQEHVATSQPKTMGQISLGLTLRLWDVYLVEGOALM
PTRIAFKVQKRLTKTSRSGPMARFCNRFVDTWARDEDTVLKHLSAKKLTTRKQD
LPPAKPEQSSASRPVPASRGKTLCKGDRAPGPPARPRPIWSASPPRAPRSST
PCPGAVREDTYPVGTQGVPSPALAQGGPQGSWRFLONSMPRLPTDL DVEGPMFRHY
DFROS CMVRAISOEDQLAPCWAQAEHPAERVRSAFAPASTDS DQTFPRARDEQPCAPT
SGPCLCGHLBESSQFP PGP"

ORIGIN
Query Match 92.1%; Score 1809; DB 6; Length 2647;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 5; Indels 66; Gaps 1;
Qy 1 ATGACGTGTAGAGGTTCGCGGCAAGTTGTTGGGCAAGAGCCGAGAGACATCATTTATG 60
Db 755 ATGACGTGTAGAGGTTCGCGGCAAGTTGTTGGGCAAGAGCCGAGAGACATCATTTATG 814
Qy 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTCA 120
Db 815 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGCAAGGGCCTTAAGCCTTTTCA 874
Qy 121 AGCTACAACAACACGTGCATCTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 180
Db 875 AGCTACAACAACACGTGCATCTTGGGGATTGTACATGAGACGAGCTGCTCTCTG 934
Qy 181 ACTGCGGGGAGGCGAAGCAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGGTGGAT 240
Db 935 ACTGCGGGGAGGCGAAGCAAATTCGGCGGGAGATCAGCCGAAAGCAAGTGGGTGGAT 994
Qy 241 ATGCTGGAGACTGGAGAAATACAAAAGCAGACAGAAAGCTCATAGATCGAGGTACAAG 300
Db 995 ATGCTGGAGACTGGAGAAATACAAAAGCAGACAGAAAGCTCATAGATCGAGGTACAAG 1054
Qy 301 GGAATGCCCATGAACATCCGGGCCCGGATGTGTGTCAGTCTCTGAAATGAGAAATG 360
Db 1055 GGAATGCCCATGAACATCCGGGCCCGGATGTGTGTCAGTCTCTGAAATGAGAAATG 1114
Qy 361 AAGTTGAAAACCCCGGAAGATACAGATCATGAAGAGAAAGGCAAGGTCATCTGAG 420
Db 1115 AAGTTGAAAACCCCGGAAGATACAGATCATGAAGAGAAAGGCAAGGTCATCTGAG 1174
Qy 421 CACATCCAGCCCATCGACCCGGGACGTAAGCGGGACATTAAGGAAGCATATATCTTCAGG 480
Db 1175 CACATCCAGCCCATCGACCCGGGACGTAAGCGGGACATTAAGGAAGCATATATCTTCAGG 1234
Qy 481 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCATTCCTCTGCGATATGAGAGTAT 540
Db 1235 GATCGATACGGAACCAAGCAGCGGGAATCTACTCCATTCCTCTGCGATATGAGAGTAT 1294
Qy 541 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATGCGCCGCTGTTCTCTCTAT 600
Db 1295 AACCCGAGGTGGGCTACTGACAGGACCTGAGCCACATGCGCCGCTGTTCTCTCTAT 1354
Qy 601 CTTCTGAGGAGATGATCATTTCTGGGCACTGTGTGACGTGCTGGCCAGTGAAGGCACTCC 660
Db 1355 CTTCTGAGGAGATGATCATTTCTGGGCACTGTGTGACGTGCTGGCCAGTGAAGGCACTCC 1414
Qy 661 CTGCAAGGATTTCAACAGCCCAAATGGCGGAGCCGTCACAGGGGCTCCAAGACCAACAGAG 720
Db 1415 CTGCAAGGATTTCAACAGCCCAAATGGCGGAGCCGTCACAGGGGCTCCAAGACCAACAGAG 1474
Qy 721 CATGTGTAGCCACGTCAACAACCCAAAGACCATGGGGCATCAGGACAAAGAAATCTATGT 780
Db 1475 CATGTGTAGCCACGTCAACAACCCAAAGACCATGGGGCATCA----- 1515
Qy 781 GGGCAGTGTCCCGGTAGGCTGCTCATCCGGAATGTGATTGACGGGATCTCTCGGG 840
Db 1516 -----GATCTCTCTCGGG 1528
Qy 841 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAAAGCGAACAGCGTTGATGCGGATA 900
Db 1529 CTCACCTGCGCCTGTGGACGTGTATCTGTAGAAAGCGAACAGCGTTGATGCGGATA 1588
Qy 901 ACAAGATCGCCTTTAAGTTACAGAGAAGCGCCTCAAGAGAGCTCCAGGTGGCCCG 960
Db 1589 ACAAGATCGCCTTTAAGTTACAGAGAAGCGCCTCAAGAGAGCTCCAGGTGGCCCG 1648
Qy 961 TGGGCAAGTTTTCGAAACCGGTTCTTGATACCTGGGCCAAGGATGAGGACACTGTGCTC 1020
Db 1649 TGGGCAAGTTTTCGAAACCGGTTCTTGATACCTGGGCCAAGGATGAGGACACTGTGCTC 1708

QY	1021	AAGCATCTTAGGGCCCTCTATGAAAGAACTAACAGAAAGCAGGGGGACCTGCCACCCCA	1080
Db	1709	AAGCATCTTAGGGCCCTCTATGAAAGAACTAACAGAAAGCAGGGGGACCTGCCACCCCA	1768
QY	1081	GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCCCTGTGCCGCTTCACTGGCGGAG	1140
Db	1769	GCCAAACCCGAGCAAGGGTCGTGGCATCCAGGCCCTGTGCCGCTTCACTGGCGGAG	1828
QY	1141	ACCCTCTGCAAGGGGACAGGCGCCCTCCAGGCCACAGCCCGGTTCCCGGGCC	1200
Db	1829	ACCCTCTGCAAGGGGACAGGCGCCCTCCAGGCCACAGCCCGGTTCCCGGGCC	1888
QY	1201	ATTTGGTCAGCTTCCCCGCCACGGGCACTCGTCTTTCACACCTGTCTGTGGGCT	1260
Db	1889	ATTTGGTCAGCTTCCCCGCCACGGGCACTCGTCTTTCACACCTGTCTGTGGGCT	1948
QY	1261	GTCGGGAAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTCTAG	1320
Db	1949	GTCGGGAAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTCTAG	2008
QY	1321	GGAGGACCTCAGGGTTCCTGAGATTCTGCACTGGAATCCATGCCCGCTCCCAAG	1380
Db	2009	GGAGGACCTCAGGGTTCCTGAGATTCTGCACTGGAATCCATGCCCGCTCCCAAG	2068
QY	1381	GACCTGGACGTAGAGGGCCCTTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGCTC	1440
Db	2069	GACCTGGACGTAGAGGGCCCTTGCTCCGCCATTATGATTTCAGACAGAGCTGTGGCTC	2128
QY	1441	CGTGCCATATCCCAGAGAGACCAGCTGGCCCCCTGTGGCAGGCTGAACACCTCGGAG	1500
Db	2129	CGTGCCATATCCCAGAGAGACCAGCTGGCCCCCTGTGGCAGGCTGAACACCTCGGAG	2188
QY	1501	CGGTGAGATCGGCTTTCGCTGCACCAGCACTGATTCGACCAGGCAACCCCTTCAGA	1560
Db	2189	CGGTGAGATCGGCTTTCGCTGCACCAGCACTGATTCGACCAGGCAACCCCTTCAGA	2248
QY	1561	GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTGCCATTG	1620
Db	2249	GCTAGGAGCAACAGCCGTGTGCTCCCACTCAGGGCCTTGCTCTGCGGCTGCCATTG	2308
QY	1621	GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCACTGGGCCAGGGCTCATGCTGAT	1680
Db	2309	GAAAGTCTCAGTTCCTCCAGGCTTCTAGAAGCACTGGGCCAGGGCTCATGCTGAT	2368
QY	1681	AATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGCTCTGTTTATTTTGGTTAA	1740
Db	2369	AATTTCCCTAGGCTTAACAACCCAGCAAGCTTCGCTCTGTTTATTTTGGTTAA	2428
QY	1741	CTTATGAAATGTATTAAGAAAGAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACACC	1800
Db	2429	CTTATGAAATGTATTAAGAAAGAGTGCAAGCTCGAGAGAGATTCAAGATGGAACACACC	2488
QY	1801	AGACCCAGATCACAAAGCCCAACCATGCCCCAGGCCCTCCAGCACCCCAAGCCACGAC	1860
Db	2489	AGACCCAGATCACAAAGCCCAACCATGCCCCAGGCCCTCCAGCACCCCAAGCCACGAC	2548
QY	1861	CATCGTTCTGAATCTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCAATGAAG	1920
Db	2549	CATCGTTCTGAATCTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCAATGAAG	2608
QY	1921	ATAACTACCTTCAAGTTTGAATAATGTTCCTGTG	1959
Db	2609	ATAACCACTTCAATGTTTGAATAATGTTCCTGTG	2647

RESULT	12
AX775949	
LOCUS	AX775949
DEFINITION	Sequence 219 from Patent WO03048202.
ACCESSION	AX775949
VERSION	AX775949.1 GI:32693667
KEYWORDS	.
	2647 bp mRNA linear PAT 14-JUL-2003

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source	CDS	ORIGIN
Homo sapiens (human)	Homo sapiens	1	Matsuda, A. and Muramatsu, S.	NF-kB activating gene	Patent: WO 03048202-A 219 12-JUN-2003;	Asahi Kasei Kabushiki Kaisha (JP)	location/Qualifiers	1. 2647	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.							/organism="Homo sapiens"		
							/mol_type="mRNA"		
							/db_xref="taxon:9606"		
							755. .2338		
							/note="unnamed protein product"		
							/codon_start=1		
							/protein_id="CAB11674.1"		
							/db_xref="GI:32693668"		
							/translation="MDVEVAVGSMWAQEREDIMKYEKHRAGLPEDKGPKPFSYNNVDHLGIHETELPPLTARBAKQIRREISRSKSWMDLGMKEYKSSRKLIDRAYKGM		
							PMNIRGPMWSVLNTEBEMKLNKNGRYQIMKEKGRSSSEHQRIDRVSGLTRKHIFER		
							DRYGTQKQRELIHLLAYEYENPEVGYCRDLSHIALFLYLPEBDAFWALVQLASER		
							HSLQGFHSPNGVTQGLQDQEHVAVTSQPKTMGHQISLGLTLRLMDVYLVEGEALM		
							PITRIAFKVOQKRLTKTSRCGPMARFCNRFVDTWARDEDVTLKHLRASMKLTKQGD		
							LQPPAKPEQSSASRPVPASRGKTLCKDRQAPPGPARFPRPIWSASPRAFRSST		
							PCPGAVREDITYPVGTQGVPSPALAOGGPGQSWRFLQMSMRPLPTDLVEGFWERHY		
							DFRQSCWRAISQEDQLAPCWMQAEHPABRVRSAFAAPSTSDQGTPEFRARDEQPCAPT		
							SGPCLGLHLESSQFPFPG"		
Query Match	92.0%;	Score 1807.4;	DB 6;	Length 2647;					
Best Local Similarity	96.3%;	Pred. No. 0;							
Matches 1887;	Conservative	0;	Mismatches	6;	Indels	66;	Gaps	1;	
1	ATGAGCGTGTAGAGGTC	CGCGGCACTTGTGGGCA	CAAGAAGGCGGCTTAAG	CGCATATTATG	60				
755	ATGACGTGTAGAGGTC	CGCGGCACTTGTGGGCA	CAAGAAGGCGGCTTAAG	CGCATATTATG	814				
61	AAATACGAAAAAGGACA	CGAGCTGGGCTGCCAGAG	CAAGGCGCTTAAGCCTTT	TGCA	120				
815	AAATACGAAAAAGGACA	CGAGCTGGGCTGCCAGAG	CAAGGCGCTTAAGCCTTT	TGCA	874				
121	AGCTACAAACAACGTC	ATTTGGGATTTGTACATG	AGACGAGCTGCCTCTTG	180					
875	AGCTACAAACAACGTC	ATTTGGGATTTGTACATG	AGACGAGCTGCCTCTTG	934					
181	ACTGCGCGGAGGCGA	AGCAATTCCGCGGAGATC	AGCCGAAGAAGCAAGTGG	TGAT	240				
935	ACTGCGCGGAGGCGA	AGCAATTCCGCGGAGATC	AGCCGAAGAAGCAAGTGG	TGAT	994				
241	ATGCTGGGAGACTGGG	AGAAATACAAAAGCAGAG	AAAGCTCATAGATCGAGCG	TACAAG	300				
995	ATGCTGGGAGACTGGG	AGAAATACAAAAGCAGAG	AAAGCTCATAGATCGAGCG	TACAAG	1054				
301	GGAATGCCCATGAACA	TCCGGGGCCCGATGTG	GTCACTCTCTGAACATTG	AGAAATG	360				
1055	GGAATGCCCATGAACA	TCCGGGGCCCGATGTG	GTCACTCTCTGAACATTG	AGAAATG	1114				
361	AAAGTGA AAAAACCC	CGGAAGATACCAAGATC	ATGAAGAGAAAGGCAAG	AGTCACTGAG	420				
1115	AAAGTGA AAAAACCC	CGGAAGATACCAAGATC	ATGAAGAGAAAGGCAAG	AGTCACTGAG	1174				
421	CACATCCAGCGCATCG	ACCGGAGCGTAAGCGG	AGACATTAAGAGACATAT	ATTCTTCAGG	480				
1175	CACATCCAGCGCATCG	ACCGGAGCGTAAGCGG	AGACATTAAGAGACATAT	ATTCTTCAGG	1234				
481	GATGATACGGAACCA	AGCAGCGGAACTACTCC	ACATCTCTGGCATATGAG	AGATAT	540				
1235	GATGATACGGAACCA	AGCAGCGGAACTACTCC	ACATCTCTGGCATATGAG	AGATAT	1294				
541	AAACCCGAGGTGGCTA	CTCAGAGGAGCTGAGCC	ACATCGCCGCTTGTTCT	CTCTAT	600				

Db 1295 AACCCGGAGGTGGGCTACTGACAGGAACTGAGCCACATCGCGCCTGTCTCTCTAT 1354
Qy 601 CTTCTGAGAGGATGATTCATTTCTGGGACTGTGACGCTGCTGGCCAGTGAGAGGACTCC 660
Db 1355 CTTCTGAGAGGATGATTCATTTCTGGGACTGTGACGCTGCTGGCCAGTGAGAGGACTCC 1414
Qy 661 CTGCAAGGATTTCAACAGCCCAATGCGGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 720
Db 1415 CTGCAAGGATTTCAACAGCCCAATGCGGGAGCCGTCCAGGGGCTCCAAAGCAACAGAG 1474
Qy 721 CATGTGTAGCCACGTCACAACCCAGACCATGGGGCATCAGGACAAAGAAATCTATGT 780
Db 1475 CATGTGTAGCCACGTCACAACCCAGACCATGGGGCATCA----- 1515
Qy 781 GGGCAGTGTCCCCGTTAGCGTGCCTCATCCGATATTGATGACGGATCTCTCGGG 840
Db 1516 -----GATCTCTCTCGGG 1528
Qy 841 CTCACCCGTGGCCTGTGGAAGGTATCTGTAGAAAGGCGAAACAGCGTTGATGCCGATA 900
Db 1529 CTCACCCGTGGCCTGTGGAAGGTATCTGTAGAAAGGCGAAACAGCGTTGATGCCGATA 1588
Qy 901 ACAAGATCGCCTTTAAGGTTACAGAGAAAGCGCCTCAAGAACGCTCCAGGTGCGCCG 960
Db 1589 ACAAGATCGCCTTTAAGGTTACAGAGAAAGCGCCTCAAGAACGCTCCAGGTGCGCCG 1648
Qy 961 TGGGCACGTTTTTGCAACCGGTTGTTGATACCTGGGCGCAGGGATGAGGACACTGTGCTC 1020
Db 1649 TGGGCACGTTTTTGCAACCGGTTGTTGATACCTGGGCGCAGGGATGAGGACACTGTGCTC 1708
Qy 1021 AAGCATCTTAGGGCCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
Db 1709 AAGCATCTTAGGGCCTCTATGAAGAACTAAACAAGAAAGCAGGGGACCTGCCACCCCA 1768
Qy 1081 GCCAAACCCGAGCAAGGGTGTGCGGATCCAGGCGCTGTGCGGCTTCAAGTGCGGGAAG 1140
Db 1769 GCCAAACCCGAGCAAGGGTGTGCGGATCCAGGCGCTGTGCGGCTTCAAGTGCGGGAAG 1828
Qy 1141 ACCCTCTGCAAGGGGGAGCAGGAGCGCCCTCCAGGCCACAGCCCGTTCCCGCGGCC 1200
Db 1829 ACCCTCTGCAAGGGGGAGCAGGAGCGCCCTCCAGGCCACAGCCCGTTCCCGCGGCC 1888
Qy 1201 ATTTGGTCACTTCCCCGCCACGGGCACTCGTCTTCCACACCCTGTCTGTGGGGCT 1260
Db 1889 ATTTGGTCACTTCCCCGCCACGGGCACTCGTCTTCCACACCCTGTCTGTGGGGCT 1948
Qy 1261 GTCCGGGAAGACACCTACCCCTGTGGGCACTCAGGGTGTGCCAGGCCCGGCTGTGCTAG 1320
Db 1949 GTCCGGGAAGACACCTACCCCTGTGGGCACTCAGGGTGTGCCAGGCCCGGCTGTGCTAG 2008
Qy 1321 GGAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAAGTCCATGCCCCGCTCCCAAG 1380
Db 2009 GGAGACCTCAGGGTCTCTGAGATTCTCTGAGTGAAGTCCATGCCCCGCTCCCAAG 2068
Qy 1381 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTCAAGACAGAGCTGTGGGTC 1440
Db 2069 GACCTGACGTAGAGGGCCCTTGTTCCGCCATTATGATTCAAGACAGAGCTGTGGGTC 2128
Qy 1441 CGTGCCATATCCCAAGAGGACCAAGTGGCCCCCTGTGGCAAGGCTGAACACCCCTGCGAG 1500
Db 2129 CGTGCCATATCCCAAGAGGACCAAGTGGCCCCCTGTGGCAAGGCTGAACACCCCTGCGAG 2188
Qy 1501 CGGGTGAATCGGCTTTGCTGCAAGCCAGCACTGATTCCGACCAAGGCAACCCCTTCAGA 1560
Db 2189 CGGGTGAATCGGCTTTGCTGCAAGCCAGCACTGATTCCGACCAAGGCAACCCCTTCAGA 2248
Qy 1561 GCTAGGAGCAACAGCCGTTGTCTCCAAGGAGGCTTGTGCGGCTTCACCTG 1620
Db 2249 GCTAGGAGCAACAGCCGTTGTCTCCAAGGAGGCTTGTGCGGCTTCACCTG 2308
Qy 1621 GAAAGTTCTAGTCCCTCAGGCTTCTAGAAGCATGTGGGCAAGGCTCATGGCTGAT 1680

Db 2309 GAAAGTTCTAGTCCCTCCAGGCTTCTAGAAGCATCTGGGCCAGGGCTCATGCTGAT 2368
Qy 1681 AATTTCCCTTAGGCTTAAACAACCCCAAGCAAGCTTCCGCTCCTGTTTATTTTGGTTAA 1740
Db 2369 AATTTCCCTTAGGCTTAAACAACCCCAAGCAAGCTTCCGCTCCTGTTTATTTTGGTTAA 2428
Qy 1741 CTTATGAAAAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 1800
Db 2429 CTTATGAAAAATGTATTAAAGAAAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACC 2488
Qy 1801 AGACCCAGATCACAAAGCCAAACCATGCCCCCAGCCCTCCAGACACCCCAAGCCACGAC 1860
Db 2489 AGACCCAGATCACAAAGCCAAACCATGCCCCCAGCCCTCCAGACACCCCAAGCCACGAC 2548
Qy 1861 CATGTTCTGAATTCTGACGACACCGTGAGCCCTGCTTGTACTTTAAACTCATGGAAG 1920
Db 2549 CATGTTCTGAATTCTGACGACACCGTGAGCCCTGCTTGTACTTTAAACTCATGGAAG 2608
Qy 1921 ATAACTAACCTTCAAGTTTGAATAATATGTTTCTGTTG 1959
Db 2609 ATAACTAACCTTCAAGTTTGAATAATATGTTTCTGTTG 2647

RESULT 13
AX833552 2647 bp DNA linear PAT 15-DEC-2003
LOCUS Sequence 676 from Patent EP1347046.
DEFINITION AX833552
ACCESSION AX833552
VERSION AX833552.1 GI:39919687
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE 1
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Makamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masubo,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 676 24-SEP-2003;
Research Association for Biotechnology (JP)
FEATURES
source location/Qualifiers
1..2647
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 92.0%; Score 1807.4; DB 6; Length 2647;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

Qy 1 ATGACGTGTTAGAGTTCGCGGCAAGTTGGTGGCAACAAGCGAGACATCATTTATG 60
Db 755 ATGACGTGTTAGAGTTCGCGGCAAGTTGGTGGCAACAAGCGAGACATCATTTATG 814
Qy 61 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 120
Db 815 AAATACGAAAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCTTAAGCCTTTTGA 874
Qy 121 AGCTACAACAACAAGTGCATTTTGGGATTGTACATGAGACGAGCTGCTCCTCTG 180
Db 875 AGCTACAACAACAAGTGCATTTTGGGATTGTACATGAGACGAGCTGCTCCTCTG 934
Qy 181 ACTGCGGGAGGCGAAGCAATTGCGGAGATCAGCCGAAAGCAAGTGGTGAT 240
Db 935 ACTGCGGGAGGCGAAGCAATTGCGGAGATCAGCCGAAAGCAAGTGGTGAT 994
Qy 241 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCTACAAG 300
Db 995 ATGCTGGAGACTGGGAGAAATACAAAGCAGCAGAAAGCTCATAGATCGAGCTACAAG 1054

Qy 301 GGAATGCCATGAACATCCGGGCCCCGATGTGTCAATCTCTGAAATTGAGAAATG 360
 Db 1055 GGAATGCCATGAACATCCGGGCCCCGATGTGTCAATCTCTGAAACTGAGGAAATG 1114
 Qy 361 AAGTTGAAAAAACCCTGGAGATACAGATCATGAGAGAAAGGCAAGAGTCACTGAG 420
 Db 1115 AAGTTGAAAAAACCCTGGAGATACAGATCATGAGAGAAAGGCAAGAGTCACTGAG 1174
 Qy 421 CACATCAGCGCATCGACCCGGGAGCTAAGCGGACATTAAGAAAGCATATATTCTCAGG 480
 Db 1175 CACATCAGCGCATCGACCCGGGAGCTAAGCGGACATTAAGAAAGCATATATTCTCAGG 1234
 Qy 481 GATCGATACGGAACCAAGCAGCGGGAATACTCCACATCCTCTGGCATATGAGAGTAT 540
 Db 1235 GATCGATACGGAACCAAGCAGCGGGAATACTCCACATCCTCTGGCATATGAGAGTAT 1294
 Qy 541 AACCCTGAGGTGGGCTACTGACAGGGAAGCTGAGCCACATCGCCCTTGTCTCTCTAT 600
 Db 1295 AACCCTGAGGTGGGCTACTGACAGGGAAGCTGAGCCACATCGCCCTTGTCTCTCTAT 1354
 Qy 601 CTTCTGAGGAGGATGCAATTCTGGGCACTGGTGACGTGCTGGCCAGTGAGAGCACTCC 660
 Db 1355 CTTCTGAGGAGGATGCAATTCTGGGCACTGGTGACGTGCTGGCCAGTGAGAGCACTCC 1414
 Qy 661 CTGACAGGATTTTCAAGCCCAAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAG 720
 Db 1415 CTGACAGGATTTTCAAGCCCAAATGGCGGAGCCGTCCAGGGGCTCCAAGACCAAGAG 1474
 Qy 721 CATGTGTAGCCACGTCACAACCCAAGACCATGGGGCATCAGACAAGAAAGATCTATGT 780
 Db 1475 CATGTGTAGCCACGTCACAACCCAAGACCATGGGGCATCA----- 1515
 Qy 781 GGGCAGTGTCCCCGTTAGGCTCCTCATCCGATATTTGACGGGATCTCTCGGG 840
 Db 1516 -----GATCTCTCTCGGG 1528
 Qy 841 CTCACCTTGGCGCTGTGGACGTATCTGTAGAAAGCGCAAGCGGTGATGCCGATA 900
 Db 1529 CTCACCTTGGCGCTGTGGACGTATCTGTAGAAAGCGCAAGCGGTGATGCCGATA 1588
 Qy 901 ACAAGATCGCCTTTAAGTTAGCAGAGAGCGCCTCAAGAAAGCGTCCAGGTGCGCCG 960
 Db 1589 ACAAGATCGCCTTTAAGTTAGCAGAGAGCGCCTCAAGAAAGCGTCCAGGTGCGCCG 1648
 Qy 961 TGGGCACTTTTGGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTCTC 1020
 Db 1649 TGGGCACTTTTGGCAACCGGTTCTGTGATACCTGGGCCAGGGATGAGGACACTGTCTC 1708
 Qy 1021 AAGCATTTAGGGCCTTATGAAGAAATAACAAGAAAGCAGGGGACCTGCCACCCCA 1080
 Db 1709 AAGCATTTAGGGCCTTATGAAGAAATAACAAGAAAGCAGGGGACCTGCCACCCCA 1768
 Qy 1081 GCCAAACCCGAGCAAGGGTCTGTGGCATCCAGGCTGTGCCGCTTACGTTGGCGGGAAG 1140
 Db 1769 GCCAAACCCGAGCAAGGGTCTGTGGCATCCAGGCTGTGCCGCTTACGTTGGCGGGAAG 1828
 Qy 1141 ACCCTCTGCAAGGGGAGCAGGAGGCCCCCTCAGGCCCCACAGCCCGGTTCCGCGGCC 1200
 Db 1829 ACCCTCTGCAAGGGGAGCAGGAGGCCCCCTCAGGCCCCACAGCCCGGTTCCGCGGCC 1888
 Qy 1201 ATTGGTCAAGTTCCTCCGCGCAAGGGACCTGTTCTTCAACACCTGTCTGTGGGGCT 1260
 Db 1889 ATTGGTCAAGTTCCTCCGCGCAAGGGACCTGTTCTTCAACACCTGTCTGTGGGGCT 1948
 Qy 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTCCCAAGCCCGGCTGGCTCAG 1320
 Db 1949 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTCCCAAGCCCGGCTGGCTCAG 2008
 Qy 1321 GGAGACCTCAGGGTCTCTGGAAGATTCCTGCAAGTGAATCCATGCCCGGCTCCCAAG 1380
 Db 2009 GGAGACCTCAGGGTCTCTGGAAGATTCCTGCAAGTGAATCCATGCCCGGCTCCCAAG 2068

Qy 1381 GACCTGACGTAGAGGGCCCTTGGTCCGCCATTAATTCAGACAGAGCTGCGGTC 1440
 Db 2069 GACCTGACGTAGAGGGCCCTTGGTCCGCCATTAATTCAGACAGAGCTGCGGTC 2128
 Qy 1441 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCCGAG 1500
 Db 2129 CGTGCCATATCCAGAGAGACCAAGCTGGCCCCCTGCTGGCAGGCTGAACACCTGCCGAG 2188
 Qy 1501 CGGTTGATCGGCTTTCGCTGCACCCAGCAGTGAATTCGACAGGGGCAACCCCTTCAGA 1560
 Db 2189 CGGTTGATCGGCTTTCGCTGCACCCAGCAGTGAATTCGACAGGGGCAACCCCTTCAGA 2248
 Qy 1561 GCTAGGAGCAACAGCCGCTGCTCCACCTCAGGCTGCTGCGGCTTCCACTTG 1620
 Db 2249 GCTAGGAGCAACAGCCGCTGCTCCACCTCAGGCTGCTGCGGCTTCCACTTG 2308
 Qy 1621 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCACTGGGCGCAGGCTCAAGCTGGAT 1680
 Db 2309 GAAAGTTCTCAGTTCCTCCAGGCTTCTAGAAGCACTGGGCGCAGGCTCAAGCTGGAT 2368
 Qy 1681 AATTCCCTAGGCTTAACAACCAAGCAAGCTTCGCGTCTGTTATTTTGGTTAA 1740
 Db 2369 AATTCCCTAGGCTTAACAACCAAGCAAGCTTCGCGTCTGTTATTTTGGTTAA 2428
 Qy 1741 CTTATGAAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACAACC 1800
 Db 2429 CTTATGAAAAATGTATTAAGAAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACAACC 2488
 Qy 1801 AGACCCAGATCAAAAGCCAACCATGCCAGCCCTCCAGCACCAGCCCAAGCAGAC 1860
 Db 2489 AGACCCAGATCAAAAGCCAACCATGCCAGCCCTCCAGCACCAGCCCAAGCAGAC 2548
 Qy 1861 CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 1920
 Db 2549 CATGTTCTGAATTTGACGACACCGTGAGCCTGCTTGTACTTTAACTCATGGAAG 2608
 Qy 1921 ATAACCTACCTCAGCTTTGAATAATGTTCTCTGTTG 1959
 Db 2609 ATAACCTACCTCAGCTTTGAATAATGTTCTCTGTTG 2647

RESULT 14
 AK095385
 LOCUS
 DEFINITION Homo sapiens cDNA FLJ38066 fis, clone CTONG2015330, moderately similar to H.sapiens mRNA for tre oncogene.
 ACCESSION AK095385
 VERSION AK095385.1 GI:21754631
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Odayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,

Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoch,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL PUBMED 14702039

TITLE

REFERENCE

AUTHORS

2

Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hirao,M., Shimizu,F., Wakebe,H., Ono,T., Hishigaki,H., Watanabe,T., Ozaki,K., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T. NEDO human cDNA sequencing project Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

3 (bases 1 to 2647)

Isogai,T. and Yamamoto,J. Direct Submission

Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team) ; 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) ; cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Source

1..2647

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CTONG2015330"

/tissue_type="tongue, tumor tissue"

/clone_fib="CTONG2"

/note="cloning vector: pME18FL3"

ORIGIN

Query Match 92.0%; Score 1807.4; DB 8; Length 2647;

Best Local Similarity 96.3%; Pred. No. 0;

Matches 1887; Conservative 0; Mismatches 6; Indels 66; Gaps 1;

QY 1 ATGACGCTGTAGAGTCCGCGGCACTTGTTGGCACAAGCCGAGAGCATCATATG 60

DB 755 ATGACGCTGTAGAGTCCGCGGCACTTGTTGGCACAAGCCGAGAGCATCATATG 814

QY 61 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGGACAAGGGCCTAAGCCTTTTGA 120

DB 815 AAATACGAAAAGGACACCGAGCTGGGCTGCCAGAGGACAAGGGCCTTAAAGCCTTTTGA 874

QY 121 AGCTACAACAACGTCGATCATTTGGGATTGTACATGAGACGGAGCTGCCTCTCTG 180

DB 875 AGCTACAACAACGTCGATCATTTGGGATTGTACATGAGACGGAGCTGCCTCTCTG 934

QY 181 ACTGCGGAGGCGAAGCAATTGCGCGGAGATCAGCCGAAAGACAAGTGGGTGAT 240

DB 935 ACTGCGGAGGCGAAGCAATTGCGCGGAGATCAGCCGAAAGACAAGTGGGTGAT 994

QY 241 ATGCTGGAGACTGGGAGAAATACAAGACGAGAAAGCTCATAGTCGAGCTTACAAG 300

DB 995 ATGCTGGGAGACTGGGAGAAATACAAAAGCAGACGAAAAGCTCATAGATCGAGCGTACAAG 1054

QY 301 GGAATGCCCATGAACATCCGGGGCCGATGTGTACGTCTCTGAACATTGAGGAATG 360

DB 1055 GGAATGCCCATGAACATCCGGGGCCGATGTGTACGTCTCTGAACATTGAGGAATG 1114

QY 361 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGGCAAGGTCATCTGAG 420

DB 1115 AAGTTGAAAAACCCCGAAGATACAGATCATGAAGAGGCAAGGTCATCTGAG 1174

QY 421 CACATCCAGCGCATCGACCGGAGCTAAGCGGGACATTAAGAGCATATATTCTTCAG 480

DB 1175 CACATCCAGCGCATCGACCGGAGCTAAGCGGGACATTAAGAGCATATATTCTTCAG 1234

QY 481 GATGATACGGAAACCAAGCAGCGGGAATCTCCACATCTCTCTGCAATGAGAGTAT 540

DB 1235 GATGATACGGAAACCAAGCAGCGGGAATCTCCACATCTCTCTGCAATGAGAGTAT 1294

QY 541 AACCCGAGGTGGGCTACTGAGGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 600

DB 1295 AACCCGAGGTGGGCTACTGAGGGACCTGAGCCACATCGCCGCTGTCTCTCTAT 1354

QY 601 CTTCCTGAGGAGATGATCTGCGGCACTGTGCACTGTGCGCAGTGAGAGGCACTCC 660

DB 1355 CTTCCTGAGGAGATGATCTGCGGCACTGTGCACTGTGCGCAGTGAGAGGCACTCC 1414

QY 661 CTGAGGAGATTTACAGCCCAATGGCGGAGCCGTCAGGGGCTCCAAGCAACAGAG 720

DB 1415 CTGAGGAGATTTACAGCCCAATGGCGGAGCCGTCAGGGGCTCCAAGCAACAGAG 1474

QY 721 CATGTGTAGCCACGTCACAAACCAAGACCATGGGCGATCAGACAAGAAAGATCTATGT 780

DB 1475 CATGTGTAGCCACGTCACAAACCAAGACCATGGGCGATCA----- 1515

QY 781 GGGCAGTGTCCCGTTAGGCTGCCCTCATCCGATATTGATTGACGGGATCTCTCGGG 840

DB 1516 -----GATCTCTCGGG 1528

QY 841 CTCACCTGCGCCTGTGGACGTGTATCTGTGAGAAAGCGAAACAGGCGTTGATGCCGATA 900

DB 1529 CTCACCTGCGCCTGTGGACGTGTATCTGTGAGAAAGCGAAACAGGCGTTGATGCCGATA 1588

QY 901 ACAAGATCGCCTTTAAGGTTACAGAGAAGCGCTCACGAAGACGTCAGGTGCGCCG 960

DB 1589 ACAAGATCGCCTTTAAGGTTACAGAGAAGCGCTCACGAAGACGTCAGGTGCGCCG 1648

QY 961 TGGGCACTTTTTCGAACCGGTTGTTGATACCTGGGCCAGGGATGAGAGCACTGTGCTC 1020

DB 1649 TGGGCACTTTTTCGAACCGGTTGTTGATACCTGGGCCAGGGATGAGAGCACTGTGCTC 1708

QY 1021 AAGCATCTTAGGGCCTCTATGAAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCCA 1080

DB 1709 AAGCATCTTAGGGCCTCTATGAAGAACTAACAAAGAAAGAGGGGACCTGCCACCCCCA 1768

QY 1081 GCCAAACCCGAGCAAGGGTGTGGCATCCAGGCGCTGTGCGGCTTCAAGTGGCGGGAG 1140

DB 1769 GCCAAACCCGAGCAAGGGTGTGGCATCCAGGCGCTGTGCGGCTTCAAGTGGCGGGAG 1828

QY 1141 ACCCTGTCAAGGGGACAGGCAAGCCCTCCAGGCCACCAAGCCCGGTTCCCGGGCCC 1200

DB 1829 ACCCTGTCAAGGGGACAGGCAAGCCCTCCAGGCCACCAAGCCCGGTTCCCGGGCCC 1888

QY 1201 ATTTGGTCAGCTTCCCGGCAAGGGCACTGTTCTTCCACACCTGTCTGTGGGGCT 1260

DB 1889 ATTTGGTCAGCTTCCCGGCAAGGGCACTGTTCTTCCACACCTGTCTGTGGGGCT 1948

QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAAGCCCGGCTGCTCAG 1320

DB 1949 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAAGCCCGGCTGCTCAG 2008

QY 1321 GGAGACCTCAGGGTCTGTGAGATTCTGTGAGTGAATCCATGCCCCGCTTCCCAAG 1380

DB 2009 GGAGACCTCAGGGTCTGTGAGATTCTGTGAGTGAATCCATGCCCCGCTTCCCAAG 2068

QY	1381	GACCTGGACGTAGAGGGCCCTTG	GGTTCGGCCATTATGATTTGACA	CAGAGCTGCTGGCTC	1440
Db	2069	GACCTGGACGTAGAGGGCCCTTG	GGTTCGGCCATTATGATTTGACA	CAGAGCTGCTGGCTC	2128
QY	1441	CGTGCCATATCCCAAGAGGACCA	GTGGCCCCCTGCTGGCAGGCTG	AACACCCCTGGGAG	1500
Db	2129	CGTGCCATATCCCAAGAGGACCA	GTGGCCCCCTGCTGGCAGGCTG	AACACCCCTGGGAG	2188
QY	1501	CGGGTGAGATCGGCTTTCCGTG	CACCCAGCACTGATTTCCGACC	AGGGCACCCCTTCAGA	1560
Db	2189	CGGGTGAGATCGGCTTTCCGTG	CACCCAGCACTGATTTCCGACC	AGGGCACCCCTTCAGA	2248
QY	1561	GCTAGGGACGAACGCCGTGTGCT	CCCACTCAGGGGCTTGGCTCTG	CGGGCTCCACTTG	1620
Db	2249	GCTAGGGACGAACGCCGTGTGCT	CCCACTCAGGGGCTTGGCTCTG	CGGGCTCCACTTG	2308
QY	1621	GAAAGTTCTCAGTTCCTCCAGG	CTTCTAGAAGCATCTGGGCAAG	GGCTCATGGCTGAT	1680
Db	2309	GAAAGTTCTCAGTTCCTCCAGG	CTTCTAGAAGCATCTGGGCAAG	GGCTCATGGCTGAT	2368
QY	1681	AATTTCCCTAGGCTTAACAAC	CCAAGCAAGCTTCGCGTCCGTT	TATTTTGGTTAA	1740
Db	2369	AATTTCCCTAGGCTTAACAAC	CCAAGCAAGCTTCGCGTCCGTT	TATTTTGGTTAA	2428
QY	1741	CTTATGAAATGTATTAAGAA	AGAGTGCAAGCTCGAGAGATTC	AGAGATGGAACACCC	1800
Db	2429	CTTATGAAATGTATTAAGAA	AGAGTGCAAGCTCGAGAGATTC	AGAGATGGAACACCC	2488
QY	1801	AGACCCAGATCACAAAGCCAAC	CATGCCAGGCCCTCCAGCAC	CCCCCAGGCCCAAGAC	1860
Db	2489	AGACCCAGATCACAAAGCCAAC	CATGCCAGGCCCTCCAGCAC	CCCCCAGGCCCAAGAC	2548
QY	1861	CATCGTTCTGAATTTGACGAC	CACCGTGAGCCTGCCCTTGTACT	TTAACTCATGGAAG	1920
Db	2549	CATCGTTCTGAATTTGACGAC	CACCGTGAGCCTGCCCTTGTACT	TTAACTCATGGAAG	2608
QY	1921	ATAACTACCTTCACGTTTGA	ATAATTAATGTTTCCCTGTTG	1959	
Db	2609	ATAACTACCTTCACGTTTGA	ATAATTAATGTTTCCCTGTTG	2647	

RESULT 15									
BC058890									
LOCUS									
DEFINITION	BC058890	2111 bp	mRNA	linear	PRI 13-JUN-2005				
	Homo sapiens hypothetical protein LOC440452, mRNA (cDNA clone MGC:64921 IMAGE:5744726), complete cds.								
ACCESSION	BC058890								
VERSION	BC058890.1	GI:37590617							
KEYWORDS	MGC.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.								
REFERENCE	1 (bases 1 to 2111)								
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Boufard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,								

	CONSRMT	Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
	TITLE	Mammalian Gene Collection Program Team
	JOURNAL	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
	PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
	REFERENCE	12477932
	AUTHORS	2 (bases 1 to 2111)
	CONSRMT	NIH MGC Project
	TITLE	Direct Submission
	JOURNAL	Submitted (01-OCT-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA
	REMARK	NIH-MGC project URL: http://mgc.nci.nih.gov
	COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca
		Martin Hirst, Thomas Zeng, Ryan Morin, Michelle Moksa, Johnson Pang, Diana Mah, Jing Wang, Kieth Fichter, Eric Chuah, Allen Delaney, Rob Kirkpatrick, Agnes Barcos, Sarah Barber, Mabel Brown-John, Steve S. Chand, William Chow, Ryan Babakaïf, Dave Wong, Corey Matsuo, Jaclyn Beland, Susan Gibson, Luis deRio, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim Macdonald, Mike R. Mayo, Josh Moran, Diana Palmquist, JR Santos, Duane Smalus, Jeff Stoltz, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Steven Jones, Rob Holt, Marco Marra.
FEATURES		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAC Plate: 119 Row: 1 Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.
source		Location/Qualifiers
	1..2111	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="MGC:64921 IMAGE:5744726"
		/tissue_type="Brain, adult medulla"
		/clone_id="NIH_MGC_119"
		/lab_host="DH10B"
		/note="Vector: pCMV-SPORT6"
gene		1..2111
		/gene="LOC440452"
		/db_xref="GeneID:440452"
CDS		85..1125
		/gene="LOC440452"
		/codon_start=1
		/product="hypothetical protein LOC440452"
		/protein_id="AAHS8890.1"
		/db_xref="GI:37590618"
		/db_xref="GeneID:440452"
		/translation="MDVVEVGSWMAQEREDIMKYEKHRAGLPEDKGKPFSGSYNNIDHFGQHETELPPLTAREVKQIRREISRSKSKVMKLGEWDYKNSRLIDRAYOGIPMNIRGPMMSVLNINEIKLKNGRGROYIMKEKGKRSSSEHIQOMDLVDVSTLRRIIFERDRYTKORELLIYLAYEENVEVGCRDLSHIALFLLYPEDEFAWLVOLASERSHLOGFHSPNGGTVOGLQDDQEHVAVTSLPKTMHQDKDLGGCCSLGLRILRIDGISLGTLRLMDVYLVEGEQAEMLIITKSPLRFSSRASRRRPVARGHVFTGSLMPGP NMTCSSILGPL"
ORIGIN		
Query Match	91.6%; Score 1799.4; DB 8; Length 2111;	
Best local Similarity	94.8%; Pred. No. 0;-	
Matches 1900; Conservative	0; Mismatches 61; Indels 44; Gaps 2;	
OY	1 ATGACGTGTAGAGGTGCGGGCAGTTGGTGCGCACACAAGCGAGAGCATCATATTATG 60	
DB	85 ATGACGTGTAGAGGTGCGGTGAGTTGGTGCGCACAGAAGCGAGAGCATCATATTATG 144	

QY 61 AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCCTAAGCCTTTTCGA 120
Db 145 AAATACGAAAAGGGACACCGAGCTGGGCTGCCAGAGACAAAGGGCCCTAAGCCTTTTCGA 204
QY 121 AGCTACAACAACAGCTGCATCTTTGGGGATTGTATCATGAGACGGAGCTGCCTCTCTG 180
Db 205 AGCTACAACAACAGCTGCATCTTTGGGGATTGTATCATGAGACGGAGCTGCCTCTCTG 264
QY 181 ACTGCGCGGAGGCGAAGCAAAATTCGGCGGGAGATCAGCCGAAAAGCAAGTGGGTGAT 240
Db 265 ACTGCGCGGAGGTGAAGCAAAATTCGGCGGGAGATCAGCCGAAAAGCAAGTGGGTGAAA 324
QY 241 ATGCTGGAGACTGGAGAAATACAAGCAGAGAAAGCTCATAGTCAGCGTACAAG 300
Db 325 ATGCTGGAGAAATGGGACACCTACAAAACAGAGAAAGCTCATAGTCAGCGTACCAG 384
QY 301 GGAATGCCATGAACATCCGGGGCCCGATGTGTCAGTCTCTGAAATTTGAGGAAATG 360
Db 385 GGAATTCCTATGAACATCCGGGGCCCGATGTGTCAGTCTCTGAAATTTGAGGAAATC 444
QY 361 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCACTGAG 420
Db 445 AAGTTGAAAAACCCCGGAAGATACAGATCATGAAGGAGAAAGGCAAGAGTCACTGAAA 504
QY 421 CACATCAGCGCATCGACCGGGAGCTAAGCGGGACATTAAGAAACATATATTCTTCAGG 480
Db 505 CACATCAGCAGATGACCTGAGCTTAAGTGGACATTAAGAGGCAATATATTCTTCAGG 564
QY 481 GATCGATAAGGAAACCAAGCAGCGGAACTACTCCACATCTCTGGCATATAGAGAGTAT 540
Db 565 GATCGATAAGGAAACCAAGCAGCGGAACTACTTCATCTCTGGCATATAGAGAGTAT 624
QY 541 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCGCCTTGTCTCTCTAT 600
Db 625 AACCCGAGGTGGGCTACTGCAAGGACCTGAGCCACATCGCGCCTTGTCTCTCTAT 684
QY 601 CTTCCCTGAGAGAGATGCATTTCTGGGCACTGTGCAAGTGTGGCCAGTGAAGGCACTCC 660
Db 685 CTTCCCTGAGAGAGATGCATTTCTGGGCACTGTGCAAGTGTGGCCAGTGAAGGCACTCC 744
QY 661 CTGCAAGGATTTCAACAGCCCAATAGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 720
Db 745 CTGCAAGGATTTCAACAGCCCAATAGCGGGACCGTCCAGGGGCTCCAAGACCAAGAG 804
QY 721 CATGTGTAGCCACGTCAACAACCAAGACCATGAGGCAATCAGGACAAAGATCTATGT 780
Db 805 CATGTGTAGCCACGTCACTACCAAGACCATGTGGCATCAGGACAAAGAAATCTATGT 864
QY 781 GGGCAGTGTCCCGCTAGGCTGCTCATCCGGATATTGAACGGATCTCTCGGG 840
Db 865 GGGCAGTGTTCGCTCTTAGGCTGCTCATCCGGATATTGAATGATGGATCTCTCGGG 924
QY 841 CTCACCTGCGCCTGTGGGACGTGTATCTGTGAAAGGCCAACAGGCGTTGATGCCGATA 900
Db 925 CTCACCTGCGCCTCTGGGACGTGTATCTGTGAAAGGCCAACAGGCGTTCACTGATTA 984
QY 901 ACAAGAAATCGCCTTTAAGGTTACAGAGAAAGCGCCTCACGAAGACGTCCAGGTGTGCCG 960
Db 985 ACAAAATTCGCTTTAAGGTTACAGAGAAAGCGCCTCACGAAGACGTCCAGGTGTGCCG 1043
QY 961 TGGGCAAGTTTTCGAACCGGTTCTGTGATACCTGGGCCAGGGATGAGACACTGTGCTC 1020
Db 1044 TGGGCAAGTTTTCGAACCGGTTCTGTGATGCTGGGCCAGGGATGATGACACTGTGCTC 1103
QY 1021 AAGCATCTTAAGGCTCTATGAAGAACTAACAGAAAGAGGGGAGCTGCACCCCCA 1080
Db 1104 AAGCATCTTAAGGCTCTATGAAGAACTAACAGAAAGAGGGGAGCTGCACCCCCA 1163
QY 1081 GCCAAACCCGAGCAAGGGTCTGCGCATCCAGGCTGTGCCGCTTCACTGTGCCGAAAG 1140
Db 1164 GCCAAACCCGAGCAAGGGTCTGTGCGCATCCAGGCTGTGCCAGCTTCAAGGGGCGGAAAG 1223
QY 1141 ACCCTTGCAAGGGGAGCAGGCGCCCTCCAGGCCCAAGCCCGGTTCCCGGGGCC 1200

Db 1224 ACCCTTGCAAGGGGAGCAGGCGCCCTCCAGGCCCAAGCCCGGTTCCCGGGGCC 1283
QY 1201 ATTTGGTCAAGCTTCCCGGCAAGGCAAGCTCGTTCTTCCACACCCCTGTCTGTGGGCT 1260
Db 1284 ATTTGGTCAAGCTTCCCGGCAAGGCAAGCTCGTTCTTCCACACCCCTGTCTGTGGGCT 1343
QY 1261 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320
Db 1344 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1403
QY 1321 GGAGGACCTCAGGGTTCCTGGAGATTCTGCAAGTGAATCCATGCCCCGCTCCCAAG 1380
Db 1404 GGAGGACCTCAGGGTTCCTGGAGATTCTGCAAGTGAATCCATGCCCCGCTCCCAAG 1463
QY 1381 GACCTGGAAGTGAAGGGCCCTTGGTTCGCCATTAATGATTTCAGACAGAGCTGTGGGTC 1440
Db 1464 GACCTGGAAGTGAAGGGACCTTGGTTCGCCATTAATGATTTCAGACAGAGCTGTGGGTC 1523
QY 1441 CGTGCCATATCCAGAGAGCAAGCTGGCCCCCTGCTGGCAGAGGCTGAACACCCCTGCGGAG 1500
Db 1524 CGTGCCATATCCAGAGAGCAAGCCGGCCACTGCTGGCAGAGGCTGAACACCCCTGCGGAG 1583
QY 1501 CGGGTGAATCGGCTTTCGCTGCA----- 1524
Db 1584 CGGGTGAATCGGCTTTCACCTGACACTGAGCCACAAAGTGGGATGACTTCCCGGCCCTG 1643
QY 1525 -----CCCAGCACTGATTCGACCAAGGCAAGCCCTTCAGAGCTAAGGAGCAAGCC 1577
Db 1644 CAGTGCGCCCAAGCACTGATTCGACCAAGGCAAGCCCTTCAGAGCTAAGGAGCAAGCA 1703
QY 1578 GTGTGCTCCCACTCAAGGCGCTTGCCTCTGCGGCTTCCACTTGAAAAGTTCTCAGTTCC 1637
Db 1704 GTGTGCTCCCACTCAAGGCGCTTGCCTCTGCGGCTTCCACTTGAAAAGTTCTCAGTTCC 1763
QY 1638 TCCAGCTTCTAAGAGCATCTGGGCGAGGCTCATGAGCTGATTAATTTCCCTAAGCTTAA 1697
Db 1764 TCCAGCTTCTAAGAGCATCTGGGCGAGGCTCATGAGCTGATTAATTTCCCTAAGCTTAA 1823
QY 1698 CAACCAAGCAAGCTTGCAGTCTGCTTTTATTTTGTAACTTATGAAGATGTATTA 1757
Db 1824 CAACCAAGCAAGCTTGCAGTCTGCTTTTATTTTGTAACTTATGAAGATGTATTA 1883
QY 1758 AGAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACCAAGACCCCAAGATCACAA 1817
Db 1884 AGAAGAGTGCAGCTCGAGAGAGATTCAAGATGGAACACACCAAGACCCCAAGATCACAA 1943
QY 1818 GCCAAGCATGCCAAGCCCTTCCAGCAAGCCCAAGCCCAAGCAAGCATGCTTGAATTCTG 1877
Db 1944 GCCAAGCATGCCAAGCCCTTCCAGCAAGCCCAAGCCCAAGCAAGCATGCTTGAATTCTG 2003
QY 1878 ACGAAGCTGAGAGCTGCTTTGTACTTTAACTCATGGAAGGATTAATCTTACAGTT 1937
Db 2004 ACGAAGCTGAGAGCTGCTTTGTACTTTAACTCATGGAAGGATTAATCTTACAGTT 2063
QY 1938 TTGAATAAATGTTTCTGTGAAA 1962
Db 2064 TTGAATAAATGTTTCTGTGAAA 2088

Search completed: April 5, 2006, 19:03:48
Job time : 9950 secs